

**PENDEKATAN QUASI-EKSPERIMENTAL & KLASIFIKASI  
MENGUNAKAN GRADIENT BOOSTING DENGAN  
SHAP DALAM ANALISI PERUBAHAN NILAI  
SEKOLAH PENERIMA PROGRAM  
MAKAN BERGIZI GRATIS**

**SKRIPSI**

**DISUSUN OLEH**

**NADIRA AULIA AFRIZA**

**2209020182**



**UMSU**

Unggul | Cerdas | Terpercaya

**PROGRAM STUDI TEKNOLOGI INFORMASI  
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI  
UNIVERSITAS MUHAMMADIYAH SUMATERA UTARA**

**MEDAN**

**2026**

**PENDEKATAN QUASI-EKSPERIMENTAL & KLASIFIKASI  
MENGUNAKAN GRADIENT BOOSTING DENGAN  
SHAP DALAM ANALISI PERUBAHAN NILAI  
SEKOLAH PENERIMA PROGRAM  
MAKAN BERGIZI GRATIS**

**SKRIPSI**

**Diajukan sebagai salah satu syarat untuk memperoleh gelar Sarjana  
Komputer (S.Kom) dalam Program Studi Teknologi Informasi, pada Fakultas  
Ilmu Komputer dan Teknologi Informasi, Universitas Muhammadiyah  
Sumatera Utara**

**NADIRA AULIA AFRIZA**

**2209020182**

**PROGRAM STUDI TEKNOLOGI INFORMASI  
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI  
UNIVERSITAS MUHAMMADIYAH SUMATERA UTARA**

**MEDAN**

**2026**

## LEMBAR PENGESAHAN

### LEMBAR PENGESAHAN

Judul Skripsi : PENDEKATAN QUASI-EKSPERIMENTAL &  
KLASIFIKASI MENGGUNAKAN GRADIENT  
BOOSTING DENGAN SHAP DALAM ANALISIS  
PERUBAHAN NILAI SEKOLAH PENERIMA  
PROGRAM MAKANAN BERGIZI GRATIS

Nama Mahasiswa : NADIRA AULIA AFRIZA

NPM : 2209020182

Program Studi : TEKNOLOGI INFORMASI

Menyetujui  
Komisi Pembimbing



(Fatma Sari Hutagalung, S.Kom, M.Kom)  
NIDN. 0117019301

Ketua Program Studi



(Fatma Sari Hutagalung, S.Kom, M.Kom.)  
NIDN. 0117019301

Dekan



(Dr. Al-Khowafizmi, S.Kom., M.Kom.)  
NIDN. 0127099201

# PERNYATAAN ORISINALITAS

## PENDEKATAN QUASI-EKSPERIMENTAL & KLASIFIKASI MENGUNAKAN GRADIENT BOOSTING DENGAN SHAP DALAM ANALISIS PERUBAHAN NILAI SEKOLAH PENERIMA PROGRAM MAKAN BERGIZI GRATIS

### SKRIPSI

Saya menyatakan bahwa karya tulis ini adalah hasil karya sendiri, kecuali beberapa kutipan dan ringkasan yang masing-masing disebutkan sumbernya.

Medan, 9 April 2026

Yang membuat pernyataan



Nadira Aulia Afriza

NPM. 2209020182

**PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK  
KEPENTINGAN AKADEMIS**

Sebagai sivitas akademika Universitas Muhammadiyah Sumatera Utara, saya bertanda tangan dibawah ini:

Nama : Nadira Aulia Afriza  
NPM : 2209020182  
Program Studi : Teknologi Informasi  
Karya Ilmiah : Skripsi

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Muhammadiyah Sumatera Utara Hak Bebas Royalti Non-Eksekutif (*Non-Exclusive Royalty free Right*) atas penelitian skripsi saya yang berjudul:


**PENDEKATAN QUASI-EKSPERIMENTAL & KLASIFIKASI  
MENGUNAKAN GRADIENT BOOSTING DENGAN  
SHAP DALAM ANALISIS PERUBAHAN NILAI  
SEKOLAH PENERIMA PROGRAM  
MAKAN BERGIZI GRATIS**

Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti Non-Eksekutif ini, Universitas Muhammadiyah Sumatera Utara berhak menyimpan, mengalih media, memformat, mengelola dalam bentuk database, merawat dan mempublikasikan Skripsi saya ini tanpa meminta izin dari saya selama tetap mencantumkan nama saya sebagai penulis dan sebagai pemegang dan atau sebagai pemilik hak cipta.

Demikian pernyataan ini dibuat dengan sebenarnya

Medan, 9 April 2026

Yang membuat pernyataan

  
Nadira Aulia Afriza  
NPM. 22090020196

## RIWAYAT HIDUP

### DATA PRIBADI

Nama Lengkap : NADIRA AULIA AFRIZA  
Tempat dan Tanggal Lahir : BINJAI, 15-04-2004  
Alamat Rumah : KOTA BINJAI  
Telepon/Faks/HP : 081265257062  
E-mail : auliaafrizan@gmail.com  
Instansi Tempat Kerja : -  
Alamat Kantor : -

### DATA PENDIDIKAN

SD : SDN 020263 TAMAT: 2015  
SMP : SMP NEGERI 1 BINJAI TAMAT: 2018  
SMK : SMA NEGERI 1 BINJAI TAMAT: 2022

## KATA PENGANTAR



Alhamdulillah puji dan syukur peneliti panjatkan kehadirat Allah Subhanahu wa Ta'ala atas segala limpahan rahmat dan hidayah-Nya kepada peneliti sehingga dapat menyelesaikan perkuliahan dan penelitian skripsi ini, dengan judul “**PENDEKATAN QUASI-EKSPERIMENTAL & KLASIFIKASI MENGGUNAKAN GRADIENT BOOSTING DENGAN SHAP DALAM ANALISIS PERUBAHAN NILAI SEKOLAH PENERIMA PROGRAM MAKAN BERGIZI GRATIS**” guna memenuhi salah satu syarat dalam dalam memperoleh gelar Sarjana Komputer pada Fakultas Ilmu Komputer dan Teknologi Informasi Universitas Muhammadiyah Sumatera Utara.

Saya tentunya berterima kasih kepada berbagai pihak dalam dukungan serta doa dalam penyelesaian skripsi. Penulis juga mengucapkan terima kasih kepada:

1. Bapak Prof. Dr. Akrim, M.Pd., Rektor Universitas Muhammadiyah Sumatera Utara (UMSU)
2. Bapak Dr. Al-Khowarizmi, S.Kom., M.Kom. Dekan Fakultas Ilmu Komputer dan Teknologi Informasi (FIKTI) UMSU.
3. Ibu Dr. Firahti Rizky, S.Kom., M.Kom Wakil Dekan 1 Fakultas Ilmu Komputer dan Teknologi Informasi (FIKTI) UMSU.
4. Bapak Mhd. Basri, S.Si, M.Kom Wakil Dekan 3 Fakultas Ilmu Komputer dan Teknologi Informasi (FIKTI) UMSU.
5. Ibu Fatma Sari Hutagalung, S.Kom., M.Kom. Ketua Program Studi Teknologi Informasi dan juga sebagai pembimbing skripsi saya.
6. Bapak Okvi Nugroho, S.Kom., M.Kom. Sekretaris Program Studi Teknologi Informasi.
7. Seluruh dosen dan staf Fakultas Ilmu Komputer dan Teknologi Informasi yang telah dengan tulus memberikan ilmu pengetahuan, bimbingan, bantuan, serta pelayanan yang baik kepada penulis selama menempuh

masa perkuliahan, sehingga penulis dapat belajar, berkembang, dan menyelesaikan pendidikan ini dengan baik.

8. Kepada cinta pertama, panutan, sekaligus sosok Istimewa dalam hidup penulis. Ayahanda tercinta Budi Darmawan terima kasih atas cucuran keringat dan kerja keras yang telah engkau tukarkan menjadi sebuah nafkah demi anakmu dapat mengenyam pendidikan sampai tingkat ini, Beliau memang tidak sempat merasakan pendidikan sampai bangku perkuliahan. Seorang Ayah menjadi alasan penulis untuk bertahan sejauh ini Penulis persembahkan skripsi ini untuk nya yang selalu mendampingi anak Perempuan pertamanya, namun segala cinta, ketulusan, dan pengorbanan yang ayah berikan tetap menjadi kekuatan terbesar dalam hidup penulis. Terima kasih telah menjadi Cahaya pertama dalam hidup penulis.
9. Pintu surgaku, ibunda Ema Ratna Terimakasih sebesar-besarnya penulis berikan kepada beliau yang tidak henti hentinya memberikan kasih sayang tanpa batas dan selalu ,menusahakan apapun untuk penulis, ibu Adalah inspirasi terbesar penulis, sumber kekuatan penulis dan alasan penulis untuk tetap maju terima kasih karna selalu berjuang untuk kehidupan penulis, untuk semua doa dan dukungan berkat ibu penulis bisa berada di titik ini. Sehat selalu dan hiduplah lebih lama lagi, harus selalu ada di setiap perjalanan dan pencapaian hidup penulis ya!
10. Ucapan terima kasih juga penulis sampaikan kepada adik-adik tercinta, javier lokeswara, kirana niwasikan Dan raziq ubaid fawwaji, yang selalu memberikan dukungan, semangat, dan keceriaan di setiap proses yang penulis lalui. Terima kasih atas kehadiran kalian yang menjadi penyemangat tersendiri bagi penulis untuk terus berjuang dan

tidak menyerah dalam menyelesaikan skripsi ini. Semoga apa yang telah penulis capai dapat menjadi inspirasi dan langkah awal untukmu dalam meraih mimpi-mimpi yang lebih tinggi, serta melanjutkan perjuangan menuju masa depan yang lebih baik di banding penulis.

11. Ucapan terima kasih kepada Sahabat penulis yang juga tak kalah penting kehadirannya, Orang-orang terdekat saya Kowi Zakyro Ginting, Meyssa Ayu Amelia, Raisha Salsabila , Abdurrahman Arif, pak halim dan Salsha Utami yang telah memberikan dukungan serta bantuan tanpa kenal lelah kepada penulis. Kehadiran kalian bukan hanya sebagai sahabat, tetapi juga sebagai tempat berbagi semangat, keluh kesah, dan kekuatan di saat penulis merasa lelah. Terima kasih karena selalu menemani, mendukung, menghibur, dan berjalan bersama dalam setiap proses yang telah dilalui. Semoga persahabatan ini senantiasa terjaga, dan langkah kita semua selalu dimudahkan untuk meraih mimpi-mimpi terbaik di masa depan.

12. Terima kasih ini saya ucapkan kepada seluruh Teman-teman Kelas E yang selalu memberikan semangat dan dukungan kepada penulis.

13. Anak Perempuan pertama dan harapan pertama orang tuanya, Nadira Aulia Afriza. Ya diri sendiri Apresiasi sebesar-besarnya untuk diri sendiri yang telah berjuang diam-diam tanpa henti. Yang hari ini berhasil membuktikan bahwa dirinya mampu tumbuh dan melewati semuanya. Yang mampu berdiri sendiri dan menyelesaikan skripsi ini. Terima kasih karena tidak menyerah sampai sejauh ini. Teruslah tumbuh menjadi pribadi yang lebih baik lagi.

**PENDEKATAN QUASI-EKSPERIMENTAL & KLASIFIKASI  
MENGUNAKAN GRADIENT BOOSTING DENGAN  
SHAP DALAM ANALISIS PERUBAHAN NILAI  
SEKOLAH PERTAMA PROGRAM  
MAKAN BERGIZI GRATIS**

**ABSTRAK**

Penelitian ini bertujuan untuk menganalisis perubahan nilai siswa pada sekolah penerima program Makanan Bergizi Gratis (MBG) dengan menggunakan pendekatan quasi-experimental dan metode klasifikasi berbasis Gradient Boosting yang diinterpretasikan menggunakan SHAP (SHapley Additive exPlanations). Pendekatan quasi-experimental digunakan untuk mengevaluasi dampak program melalui perbandingan kondisi sebelum dan sesudah pelaksanaan program. Data penelitian meliputi nilai akademik siswa, tingkat kehadiran, serta faktor pendukung lainnya. Model Gradient Boosting digunakan untuk mengklasifikasi perubahan nilai siswa, sedangkan SHAP digunakan untuk menjelaskan kontribusi masing-masing variabel terhadap hasil prediksi. Hasil penelitian menunjukkan bahwa program mbg memberikan pengaruh positif terhadap peningkatan nilai siswa. Faktor yang paling berpengaruh positif terhadap peningkatan nilai siswa. Faktor yang paling berpengaruh dalam model adalah tingkat kehadiran dan kondisi ekonomi siswa.

Kata Kunci: Quasi-Ekperimental, Gradient Boosting, SHAP, Classification,

MBG

QUASI-EXPERIMENTAL APPROACH AND CLASSIFICATION  
ASSIFICATION USING GRADIENT BOOSTING WITH  
SHAP IN ANALYZING CHANGES IN ACADEMIC  
PERFORMANCE OF SCHOOLS RECEIVING  
THE FREE NUTRITIOUS MEAL  
PROGRAM

**ABSTRACT**

This study aims to analyze changes in student academic performance in schools receiving the Free Nutritious Meal (MBG) program using a quasi-experimental approach and a classification method based on Gradient Boosting with interpretation using SHAP (SHapley Additive exPlanations). The quasi-experimental approach is applied to evaluate the impact of the program by comparing conditions before and after its implementation. The data used in this study include students' academic scores, attendance rates, and other supporting factors. The Gradient Boosting model is employed to classify changes in student performance, while SHAP is used to explain the contribution of each variable to the model's predictions. The results indicate that the MBG program has a positive impact on improving student academic performance. Attendance and economic conditions are identified as the most influential factors in the model.

*Keywords: Robotics; Quasi-Experimental, Gradient Boosting, SHAP, Classification, MBG*

## DAFTAR ISI

LEMBAR PENGESAHAN .....	ii
PERNYATAAN ORISINALITAS .....	iii
PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS .....	iv
RIWAYAT HIDUP.....	v
KATA PENGANTAR .....	vi
ABSTRAK .....	ix
ABSTRACT .....	x
BAB I.....	1
PENDAHULUAN .....	1
1.1    Latar Belakang Masalah.....	1
1.2    Rumusan Masalah .....	4
1.3    Batasan Masalah.....	4
1.4    Tujuan Penelitian .....	5
1.5    Manfaat Penelitian .....	5
BAB II.....	7
LANDASAN TEORI.....	7
2.1    Program Makan Bergizi Gratis (MBG).....	7
2.1.1 Konsep dan Tujuan MBG .....	7
2.1.2 Dampak MBG terhadap gizi dan prestasi akademik.....	9
2.1.3 Studi Empiris Terkait MBG di Indonesia dan Luar Negeri .....	10
2.2    Evaluasi Quasi-Eksperimental .....	11
2.2.1 Pendekatan Quasi-Eksperimental .....	11
2.2.2 Kelebihan dan Keterbatasan Quasi-Eksperimental dalam Penelitian Pendidikan.....	15
2.3    Machine Learning untuk Analisis Pendidikan .....	16
2.3.1 Konsep Machine Learning dalam Pendidikan.....	17
2.3.2 Algoritma Gradient Boosting .....	17
2.3.3 Interpretabilitas Model dengan SHAP.....	19
2.4    Penelitian Terkait .....	22
BAB III .....	25
METODOLOGI PENELITIAN.....	25
3.1    Pendekatan Penelitian .....	25

3.2	Variabel Penelitian .....	25
3.3	Tahapan penelitian .....	26
3.4	Model dan Metode Analisis .....	27
3.4.1	Propensity Score Matching (PSM).....	27
3.4.2	Difference-in-Differences (DiD) .....	29
3.4.3	Algoritma Gradient Boosting .....	29
3.4.4	SHAP (Shapley Additive exPlanations).....	30
3.5	Teknik Pengumpulan Data.....	31
3.6	Perancangan Sistem .....	31
3.6.1	Arsitektur Sistem.....	31
3.7	Perancangan UML (Unified Modeling Language) .....	33
3.7.1	Activity Diagram .....	34
3.7.2	Use Case Diagram .....	35
3.7.3	Sequence Diagram.....	37
3.7.4	Data Flow Diagram Level 0 (Context Diagram) .....	39
3.7.5	Data Flow Diagram Level 1 .....	40
3.7.6	Data Flow Diagram Level 2 .....	42
3.7.7	Class Diagram (Rancangan Basis Data).....	44
3.8	Perancangan Interface .....	46
<b>BAB IV HASIL DAN PEMBAHASAN .....</b>		<b>48</b>
4.1	Hasil Perancangan Sistem .....	48
4.1.1	Tampilan Halaman Login.....	49
4.2	Statistik Deskriptif Data Penelitian.....	50
4.2.1	Ringkasan Statistik Deskriptif .....	50
4.2.2	Komposisi dan Karakteristik Sample .....	51
4.3	Hasil Analisis Propensity Score Matching (PSM) .....	52
4.3.1	Proses Matching dan Hasil Keseimbangan .....	53
4.4	Hasil Analisis Difference-in-Differences (DiD).....	55
4.4.1	Verifikasi Asumsi Parallel Trends.....	55
4.4.2	Estimasi Average Treatment Effect (ATE) .....	56
4.5	Hasil Klasifikasi Gradient Boosting.....	57
4.5.1	Konfigurasi dan Evaluasi Model .....	58
4.5.2	Confusion Matrix.....	58
4.6	Analisis SHAP (Shapley Additive exPlanations) .....	61
4.6.1	Global Feature Importance .....	61
4.7	Hasil Prediksi Individual dan Segmentasi Siswa .....	64

4.7.1	Fitur Prediksi Dampak Akademik Individual.....	64
4.7.2	Segmentasi Seluruh Siswa.....	65
4.8	Tampilan Kesimpulan Otomatis dan Rekomendasi.....	66
BAB V PENUTUP.....		68
5.1	Kesimpulan .....	68
5.1.1	Dampak Signifikan MBG terhadap Prestasi Akademik .....	68
5.1.2	Kebersihan Keseimbangan Kolompok melalui PSM .....	69
5.1.3	Performa Model Gradient Boosting dalam Klasifikasi Dampak ..	69
5.1.4	Faktor-Faktor yang Mempengaruhi Dampak MBG .....	69
5.1.5	Keberhasilan Impelementasi Dashboard .....	70
5.2	Saran.....	70
5.2.1	Saran untuk Implementasi Program MBG .....	70
5.2.2	Saran untuk Pengembangan Penelitian Selanjutnya.....	71
5.2.3	Saran untuk Pengembangan Dashboard .....	72
DAFTAR PUSTAKA .....		73

## DAFTAR TABEL

	<b>HALAMAN</b>
Tabel 2. 1 Ringkasan Penelitian Terdahulu .....	23
Tabel 4. 1 Statistik Deskriptif Variabel Numerik .....	50
Tabel 4. 2 Standardized Mean Difference (SMD) Sebelum dan Sesudah Matching .....	53
Tabel 4. 3 Hasil Estimasi DiD Average Treatment Effect (ATE) .....	56
Tabel 4. 4 Metrik Evaluasi Model Gradient Boosting .....	58
Tabel 4. 5 Confusion Matrix Model Gradient Boosting .....	61
Tabel 4. 6 Peringkat Global Feature Importance (SHAP) .....	62
Tabel 4. 7 Distribusi Segmentasi Siswa .....	66

## DAFTAR GAMBAR

### HALAMAN

Gambar 3. 1 Alur penelitian.....	27
Gambar 3. 2 Propensity Score.....	28
Gambar 3. 3 Difference-in-Differences (DiD).....	29
Gambar 3. 4 Proses Gradient Boosting .....	30
Gambar 3. 5 Arsitektur Sistem.....	32
Gambar 3. 6 Activity Diagram Alur Tahapan Penelitian Analisis Dampak MBG	35
Gambar 3. 7 Use Case Diagram Sistem Analisis Dampak Program MBG .....	36
Gambar 3. 8 Sequence Diagram Alur Temporal Eksekusi Analisis Dampak MBG .....	38
Gambar 3. 9 DFD Level 0-Context Diagram Sistem Analisis Dampak MBG .....	40
Gambar 3. 10 DFD Level 1-Proses Utama Sistem Analisis Dampak MBG.....	41
Gambar 3. 11 DFD Level 2 Sub-Proses Analisis Quasi-Eksperimental dan Machnie Learning .....	43
Gambar 3. 12 Class Diagram Rancangan Basis Data Sistem Analisis MBG (Normalisasi 3NF).....	45
Gambar 3. 13 Desain Halaman Dashboard Sistem .....	46
Gambar 4. 1 Tampilan Halaman Login Dashboard MBG .....	49
Gambar 4. 2 Tampilan Halaman Eksplorasi Data (EDA) Statistik Deskriptif .....	51
Gambar 4. 3 Tampilan Halaman Ikhtisar Ringkasan Statistik dan Tren DiD.....	52
Gambar 4. 4 Hasil Analisis PSM Tabel SMD dan Love Plot .....	54
Gambar 4. 5 Hasil Analisis DiD ATE Cards, Parallel Trends, dan Forest Plot....	57
Gambar 4. 6 Hasil Analisis Gradient Boosting Metrik Evaluasi dan Confusion Matrix.....	59
Gambar 4. 7 Analisis SHAP Global Feature Importance dan Beeswarm Plot .....	62
Gambar 4. 8 Tampilan Form Prediksi Dampak Akademik Individual .....	64
Gambar 4. 9 Hasil Segmentasi Siswa Distribusi Prediksi Seluruh Siswa .....	65
Gambar 4. 10 Tampilan Halaman Kesimpulan & Rekomendasi.....	67

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang Masalah

Program Makan Bergizi Gratis (MBG) merupakan kebijakan strategis pemerintah Indonesia yang bertujuan meningkatkan gizi, konsentrasi belajar, dan capaian akademik siswa. Berbagai penelitian menunjukkan bahwa MBG berdampak positif terhadap performa siswa, terutama dalam peningkatan konsentrasi dan hasil belajar, khususnya di bidang matematika (Arifin et al., 2025); (Yesi & Annur, 2025). Program ini juga terbukti meningkatkan status gizi, kehadiran, serta kesehatan fisik dan mental peserta didik (Sinaga & Syarief, 2025); (Prasetyo et al., 2025); (Nababan et al., 2024). Selain manfaat akademik, MBG turut meningkatkan motivasi dan partisipasi belajar melalui dukungan komunitas sekolah dan keterlibatan orang tua (Solihat et al., 2025); (Sarwati et al., 2025). Dari sisi tata kelola, keberhasilan program ini sangat dipengaruhi oleh komitmen politik, kesinambungan pendanaan, dan efektivitas distribusi makanan (Marnia et al., 2024); (Fitriani & Hidayat, 2024).

Permasalahan utama program Makan Bergizi Gratis (MBG) adalah belum adanya pemetaan empiris yang jelas antara pelaksanaan program dan peningkatan capaian akademik. Meski banyak laporan kebijakan mengklaim keberhasilan MBG, temuan lapangan menunjukkan hasil yang bervariasi antarwilayah dan jenjang sekolah. Beberapa studi menunjukkan peningkatan signifikan nilai akademik di Banda Aceh dan Banten pada siswa penerima MBG (Arifin et al., 2025); (Solihat et al., 2025), sementara penelitian lain menemukan efek yang

tidak konsisten akibat kendala logistik dan variasi kualitas menu (Sinaga & Syarief, 2025). Variasi ini menimbulkan pertanyaan akademik mengenai faktor kontekstual seperti kondisi sosial ekonomi, infrastruktur sekolah, dan konsistensi kebijakan (Marnia et al., 2024).

Karena hasil belajar mencerminkan kualitas pendidikan, evaluasi berbasis data diperlukan untuk memastikan dampak MBG yang positif, merata, dan berkelanjutan serta mendukung peninjauan desain dan mekanisme program.

Efektivitas program Makan Bergizi Gratis (MBG) penting dikaji karena hasil belajar mencerminkan kualitas pendidikan nasional. Jika dampak akademiknya belum konsisten, perlu peninjauan terhadap desain dan pelaksanaannya. Keberhasilan program sangat ditentukan oleh kualitas implementasi, dukungan komunitas, dan koordinasi lintas sektor (Andin et al, 2024); (Hapsari et al, 2024). Selain itu, kebijakan berbasis data dan konteks local terbukti lebih efektif dalam meningkatkan hasil belajar secara berkelanjutan. (Rizki & Darmayanti, 2025). Oleh karena itu, evaluasi empiris diperlukan untuk memastikan MBG benar-benar efektif, adil, dan berkelanjutan.

Penelitian ini menggunakan pendekatan quasi-eksperimental yang dikombinasikan dengan Gradient Boosting dan SHAP untuk mengevaluasi dampak program Makan Bergizi Gratis (MBG) terhadap capaian akademik siswa secara empiris dan berbasis bukti. Pendekatan ini dipilih karena efektif mengevaluasi kebijakan tanpa memerlukan randomisasi penuh, yang sulit diterapkan dalam program berskala nasional (Amini-Rarani et al., 2024). Sekolah penerima program dijadikan kelompok perlakuan, sedangkan sekolah serupa yang belum menerima program menjadi kelompok kontrol. Untuk menyeimbangkan

perbedaan karakteristik awal, digunakan Propensity Score Matching (PSM), kemudian dianalisis menggunakan metode Difference-in-Differences (DiD) guna mengukur perubahan capaian akademik secara objektif (Bracic et al., 2022).

Data penelitian terdiri atas tiga komponen utama, yaitu: (1) variabel hasil berupa perubahan nilai akademik siswa sebelum dan sesudah program, (2) variabel perlakuan yang menunjukkan status penerimaan MBG, serta (3) variabel kontrol seperti kondisi sosial ekonomi keluarga, tingkat kehadiran, dan kualitas sarana sekolah. Data diperoleh dari dokumen akademik sekolah, laporan dinas kesehatan, serta survei pendukung.

Dalam tahap analisis, algoritma Gradient Boosting dipilih karena memiliki keunggulan empiris dan teknis dalam mengolah data pendidikan yang kompleks. Algoritma ini terbukti memberikan akurasi prediksi tertinggi dalam memodelkan hubungan non-linear antara variabel akademik, melampaui metode tradisional seperti regresi linier dan SVM (Özkurt, 2024). Selain itu, Gradient Boosting unggul dalam mendeteksi interaksi kompleks antarvariabel dan lebih efisien dibandingkan model ensemble lain seperti Random Forest atau Neural Network dalam menangani data tidak seimbang dan ber-noise tinggi (Cortez & Silva, 2020); (Rizki & Darmayanti, 2025). Pendekatan ini dinilai paling optimal untuk menggambarkan dinamika perubahan capaian akademik siswa secara akurat dan andal.

Tahapan implementasi mencakup pra-pemrosesan data, pelatihan model dengan pembagian *training* dan *testing set*, serta evaluasi performa menggunakan metrik  $R^2$ , MAE, dan RMSE. Untuk memastikan transparansi hasil, model diintegrasikan dengan SHAP (SHapley Additive exPlanations) guna menjelaskan

kontribusi tiap variabel terhadap prediksi (Yang, Lee, & Kim, 2025). Pendekatan gabungan antara quasi-eksperimental, Gradient Boosting, dan SHAP ini tidak hanya mengukur dampak program MBG secara empiris, tetapi juga mengungkap mekanisme penyebab perubahan nilai akademik secara transparan dan dapat dipertanggungjawabkan.

## **1.2 Rumusan Masalah**

Bagaimana pendekatan Quasi-Eksperimental dan klasifikasi menggunakan gradient boosting dengan shap dalam analisis perubahan nilai sekolah penerimaan program makanan bergizi gratis

## **1.3 Batasan Masalah**

Penelitian ini memiliki batasan permasalahan yang dibahas agar dapat menyelesaikan permasalahan utama. Berikut ini batasan masalah penelitian ini:

1. Penelitian ini dibatasi pada siswa Sekolah Dasar penerima Program Makan Bergizi Gratis (MBG) yang mulai dilaksanakan pada tanggal 03 November 2025 di Kecamatan Binjai Utara, Kota Binjai.
2. Lokasi penelitian hanya mencakup sekolah penerima MBG di Kecamatan Binjai Utara, yaitu SD Yaspen Melati, SD Muhammadiyah 04, SD Negeri 027144, MIS Miftahul Jannah, SD Negeri 023906, dan SD Negeri 024760.
3. Variabel yang dianalisis terbatas pada perubahan nilai akademik siswa sebelum dan sesudah pelaksanaan Program Makan Bergizi Gratis (MBG).
4. Data penelitian menggunakan data akademik siswa dan data pendukung yang tersedia dari administrasi sekolah selama periode sebelum dan sesudah pelaksanaan program.

5. Analisis penelitian menggunakan pendekatan quasi-eksperimental (Propensity Score Matching dan Difference-in-Differences) serta algoritma Gradient Boosting dengan interpretasi SHAP untuk mengidentifikasi dampak program.

#### **1.4 Tujuan Penelitian**

Tujuan dari penelitian ini adalah:

1. Menjajaki dan membuktikan pengaruh program Makan Bergizi Gratis (MBG) terhadap perubahan capaian akademik siswa Sekolah Dasar (SD).
2. Menguraikan dan mengukur besarnya pengaruh program MBG terhadap peningkatan nilai akademik setelah dilakukan penyeimbangan karakteristik awal menggunakan quasi-eksperimental.
3. Menerapkan algoritma Gradient Boosting untuk mengklasifikasikan perubahan nilai akademik siswa penerima MBG secara akurat.
4. Menjelaskan faktor-faktor yang paling berkontribusi terhadap perubahan nilai akademik siswa penerima MBG melalui analisis interpretabilitas menggunakan SHAP (Shapley Additive Explanations).

#### **1.5 Manfaat Penelitian**

Manfaat dari penelitian ini antara lain adalah:

1. Memberikan evaluasi empiris mengenai pengaruh Program Makan Bergizi Gratis (MBG) terhadap perubahan capaian akademik siswa Sekolah Dasar.
2. Menghasilkan analisis berbasis data menggunakan pendekatan quasi-eksperimental untuk mengukur efektivitas program secara objektif dan akurat.

3. Mengoptimalkan penerapan algoritma Gradient Boosting dan SHAP dalam mengklasifikasikan serta menjelaskan faktor yang memengaruhi perubahan nilai akademik siswa.
4. Menjadi referensi dalam pengembangan kebijakan pendidikan dan program gizi sekolah yang lebih efektif, adil, dan berkelanjutan.

## **BAB II**

### **LANDASAN TEORI**

#### **2.1 Program Makan Bergizi Gratis (MBG)**

Program Makan Bergizi (MBG) merupakan bagian dari strategi peningkatan kualitas pendidikan yang berfokus pada pemenuhan kebutuhan dasar anak, khususnya dalam hal gizi. Menurut Suryani dan Putra (2022) siswa yang mengikuti program MBG menunjukkan tingkat konsentrasi dan kehadiran yang lebih tinggi dibandingkan siswa yang tidak menerima intervensi serupa. Kondisi ini disebabkan oleh peningkatan asupan energi dan zat gizi makro yang dibutuhkan otak dalam menjalankan fungsi kognitif. Penelitian tersebut memperkuat argumen bahwa keberhasilan akademik tidak hanya bergantung pada kualitas pembelajaran, tetapi juga pada kesiapan fisiologis dan psikologis siswa. Hal ini sejalan dengan pandangan Hoddinott et al. (2021) yang menegaskan bahwa intervensi gizi sejak dini berpengaruh besar terhadap perkembangan otak dan kemampuan berpikir anak, yang pada akhirnya berdampak positif pada prestasi belajar. (Saputra, Erwin, 2025).

##### **2.1.1 Konsep dan Tujuan MBG**

###### **1. Konsep MBG (Makanan Bergizi Gratis)**

Program Makan Bergizi Gratis (MBG) merupakan program intervensi pemerintah di bidang pendidikan dan kesehatan yang bertujuan menyediakan makanan dengan kandungan gizi seimbang bagi peserta didik secara gratis di lingkungan sekolah. Konsep utama MBG berfokus

pada pemenuhan kebutuhan gizi harian siswa guna mendukung pertumbuhan fisik, kesehatan, dan perkembangan kognitif.

MBG dirancang sebagai upaya preventif untuk mengurangi masalah gizi, seperti kekurangan energi, anemia, dan stunting, sekaligus sebagai strategi pendukung peningkatan kualitas pembelajaran. Dengan terpenuhinya asupan gizi yang memadai, siswa diharapkan memiliki kondisi fisik yang lebih baik, tingkat konsentrasi yang lebih tinggi, serta kesiapan belajar yang optimal selama proses pembelajaran berlangsung.

## **2. Tujuan Program Makan Bergizi Gratis (MBG)**

### **a. Tujuan Umum**

Meningkatkan status gizi dan kualitas kesehatan peserta didik guna mendukung proses pembelajaran yang efektif dan meningkatkan kualitas sumber daya manusia.

### **b. Tujuan Khusus**

1. Memenuhi kebutuhan gizi seimbang peserta didik secara merata dan berkelanjutan.
2. Mengurangi risiko kekurangan gizi dan masalah kesehatan pada anak usia sekolah.
3. Meningkatkan konsentrasi, kehadiran, dan motivasi belajar siswa.
4. Mendukung peningkatan prestasi dan hasil belajar siswa.
5. Mengurangi kesenjangan sosial dalam akses terhadap makanan bergizi di lingkungan sekolah.

6. Mendukung kebijakan pembangunan pendidikan dan kesehatan berbasis kesejahteraan anak.

### **2.1.2 Dampak MBG terhadap gizi dan prestasi akademik**

Program Makan Bergizi Gratis memberikan dampak positif terhadap pemenuhan kebutuhan gizi peserta didik. Penyediaan makanan dengan komposisi gizi seimbang membantu meningkatkan asupan energi, protein, vitamin, dan mineral yang dibutuhkan tubuh untuk pertumbuhan dan kesehatan. Program ini berperan dalam mengurangi risiko kekurangan gizi, anemia, serta masalah gizi lainnya pada anak usia sekolah. Dengan status gizi yang lebih baik, daya tahan tubuh siswa meningkat sehingga angka sakit dan ketidakhadiran di sekolah dapat berkurang.

Semakin teratur dan seimbang pola makan anak, semakin baik pula pencapaian akademik mereka. Hal ini menunjukkan bahwa gizi yang memadai tidak hanya berperan dalam pertumbuhan fisik, tetapi juga sangat penting untuk perkembangan otak dan proses belajar anak. Secara keseluruhan, hasil-hasil penelitian menegaskan bahwa pola makan sehat dan bergizi sangat penting untuk mendukung kesuksesan belajar anak. Dengan memastikan anak-anak mengonsumsi makanan yang bernutrisi, kita bisa membantu mereka menjadi lebih fokus, bersemangat, dan siap menyerap pelajaran dengan optimal di sekolah.

Makanan bergizi sangat mendukung peningkatan konsentrasi belajar, terutama pada anak-anak yang mengalami kesulitan dalam belajar. Hasil observasi Khusna Yulinda Udhiyanasari menunjukkan bahwa siswa yang tidak sarapan atau hanya mengonsumsi makanan yang tidak seimbang cenderung mengalami penurunan fokus dan mudah mengantuk. Sebaliknya, siswa yang

secara rutin mengonsumsi sarapan bergizi menunjukkan peningkatan konsentrasi serta respons yang lebih baik dalam kegiatan belajar (Udhiyanasari, 2023). Berdasarkan berbagai temuan literatur dan observasi, dapat disimpulkan bahwa pola makan yang sehat dan bergizi memiliki pengaruh signifikan terhadap kemampuan belajar anak. Asupan nutrisi yang cukup dan seimbang tidak hanya mendukung pertumbuhan fisik, tetapi juga berperan penting dalam meningkatkan fungsi otak, konsentrasi, dan daya serap siswa selama proses pembelajaran. Anak-anak yang terbiasa sarapan dan mengonsumsi makanan bergizi cenderung lebih fokus, tidak mudah mengantuk, serta memiliki prestasi akademik yang lebih baik dibandingkan dengan mereka yang kekurangan gizi. Oleh karena itu, memastikan kecukupan gizi dalam pola makan anak merupakan langkah penting untuk mendukung kesuksesan belajar dan perkembangan kognitif mereka di sekolah.

### **2.1.3 Studi Empiris Terkait MBG di Indonesia dan Luar Negeri**

#### **Studi Empiris di Indonesia**

Penelitian yang dilakukan di SDN 31 Palembang menunjukkan bahwa pelaksanaan program makanan gratis secara rutin berdampak positif terhadap konsentrasi, memori, motivasi belajar, dan prestasi akademik siswa. Studi ini menggunakan pendekatan deskriptif dan observasi nilai sebelum dan sesudah program, serta wawancara dengan guru untuk menggambarkan perubahan yang terjadi. Hasil penelitian menyatakan bahwa program ini mampu memberikan dampak akademik sosial yang signifikan apabila dievaluasi dan dijalankan secara berkelanjutan. (Al Fazri & Nabila, 2025).

#### **Studi Empiris di Luar Negeri**

Penelitian ini mengevaluasi dampak program nutrisi sekolah di sejumlah sekolah menengah pertama di Sri Lanka. Dengan desain *mixed-methods*, ditemukan bahwa program makan di sekolah berkontribusi pada penurunan absensi siswa sekitar 20% dan peningkatan nilai rata-rata ulangan sekitar 15%, serta memperbaiki konsentrasi dan keterlibatan siswa di kelas. Hasil wawancara dengan guru dan siswa mengindikasikan kehadiran yang lebih baik dan lebih sedikit masalah perilaku di sekolah dengan program gizi sekolah. (Rozi & Darlan, 2018). Review sistematis dan meta-analisis menunjukkan bahwa program makan di sekolah secara signifikan meningkatkan berat badan dan tinggi badan anak, serta meningkatkan persentase kehadiran di sekolah, yang menjadi faktor pendukung prestasi akademik. (Wang et al., 2021) Studi komparatif di negara-negara Afrika menunjukkan hubungan positif antara program makan di sekolah dengan hasil tes akademik, kehadiran, dan tingkat pendaftaran peserta didik. (Wall et al., 2022).

Penelitian yang meninjau program pemberian makanan sekolah di Ghana menunjukkan bahwa program tersebut memiliki dampak positif terhadap kehadiran siswa dan prestasi akademik, terutama di daerah pedesaan. Meskipun studi ini menggunakan pendekatan studi pustaka (*desk study*), temuan menunjukkan bahwa pengurangan kelaparan melalui program makan sekolah dapat meningkatkan keterlibatan siswa dalam pembelajaran dan menurunkan angka putus sekolah. (Appiah, 2024).

## **2.2 Evaluasi Quasi-Eksperimental**

### **2.2.1 Pendekatan Quasi-Eksperimental**

Pendekatan quasi-eksperimental merupakan metode penelitian kuantitatif

yang digunakan untuk menguji pengaruh suatu perlakuan terhadap variabel tertentu tanpa melibatkan randomisasi subjek secara penuh. Dalam pendekatan ini, peneliti membandingkan kelompok eksperimen yang menerima perlakuan dengan kelompok kontrol yang tidak menerima perlakuan, di mana kedua kelompok telah terbentuk sebelumnya, seperti kelas yang sudah ada di sekolah. Umumnya, desain yang digunakan adalah pretest–posttest, sehingga peneliti dapat mengamati perubahan kondisi sebelum dan sesudah perlakuan diberikan. Pendekatan quasi-eksperimental banyak digunakan dalam penelitian pendidikan karena dinilai mampu memberikan bukti empiris yang kuat dalam kondisi pembelajaran nyata meskipun terdapat keterbatasan dalam pengendalian variabel penelitian. (Widyasari & Rohmah, n.d.).

#### **2.2.1.1 Konsep Quasi-Eksperimental**

Pendekatan quasi-eksperimental merupakan metode penelitian kuantitatif yang digunakan untuk mengkaji hubungan sebab-akibat antara perlakuan (treatment) dan hasil (outcome) tanpa melibatkan proses pengacakan subjek penelitian secara penuh. Pendekatan ini banyak digunakan dalam penelitian pendidikan karena keterbatasan peneliti dalam melakukan randomisasi peserta didik, mengingat kelas atau kelompok belajar telah terbentuk secara alami. Oleh karena itu, quasi-eksperimental menjadi alternatif yang relevan untuk mengevaluasi efektivitas suatu program atau kebijakan pendidikan dalam kondisi nyata. Desain yang umum digunakan adalah pretest–posttest control group design, di mana terdapat kelompok eksperimen yang menerima perlakuan dan kelompok kontrol sebagai pembanding, dengan pengukuran dilakukan sebelum dan sesudah perlakuan untuk mengetahui perubahan akibat intervensi yang diberikan.

Perbandingan selisih hasil pretest dan posttest antara kedua kelompok memungkinkan peneliti menilai sejauh mana perlakuan memberikan dampak yang signifikan terhadap variabel penelitian. Meskipun tidak menggunakan randomisasi penuh, pendekatan quasi-eksperimental tetap mampu menghasilkan bukti empiris yang kuat apabila dikombinasikan dengan teknik analisis statistik lanjutan, seperti pencocokan karakteristik awal atau pengendalian kovariat, sehingga temuan penelitian dapat dinilai valid dan reliabel dalam konteks penelitian pendidikan dan kebijakan publik. (Haka et al., 2025)

#### **2.2.1.2 Propensity Score Matching (PSM)**

Propensity Score Matching (PSM) sebagai metode untuk mengestimasi dampak suatu kebijakan atau intervensi dalam kondisi penelitian non eksperimental. Penelitian ini bertujuan untuk mengurangi bias seleksi yang muncul akibat perbedaan karakteristik awal antara kelompok perlakuan dan kelompok kontrol. Melalui perhitungan propensity score berdasarkan variabel kovariat yang relevan, subjek penelitian dipasangkan sehingga kedua kelompok memiliki karakteristik awal yang seimbang. Hasil penelitian menunjukkan bahwa penerapan PSM mampu meningkatkan validitas estimasi dampak perlakuan dibandingkan analisis tanpa pencocokan, serta memberikan gambaran yang lebih akurat mengenai pengaruh kebijakan yang diteliti. Jurnal ini menegaskan bahwa PSM merupakan pendekatan yang efektif dan reliabel untuk evaluasi kebijakan dan program sosial, khususnya ketika randomisasi tidak memungkinkan dalam desain penelitian. (Melkamu et al., 2022).

Propensity Score Matching (PSM) adalah teknik statistik yang digunakan untuk mengestimasi *efek kausal* suatu perlakuan atau intervensi dalam penelitian

observasional/quasi-eksperimental dengan tidak adanya randomisasi subjek. PSM bekerja dengan menghitung skor kecenderungan (propensity score), yaitu probabilitas setiap individu menerima perlakuan berdasarkan variabel kovariat awal yang relevan (misalnya: usia, jenis kelamin, kondisi sosial ekonomi). Selanjutnya, individu pada kelompok perlakuan dipasangkan dengan individu pada kelompok kontrol yang memiliki nilai skor yang serupa, sehingga kedua kelompok menjadi lebih seimbang pada karakteristik awal mereka. Dengan demikian, perbandingan hasil antar kelompok setelah pencocokan dapat mendekati kondisi eksperimen acak (*randomized controlled trial*) dan memungkinkan pengujian efek perlakuan secara lebih valid pada data observasional. PSM banyak digunakan dalam evaluasi kebijakan publik, pendidikan, kesehatan, dan ilmu sosial lainnya untuk mengurangi bias seleksi dalam estimasi dampak intervensi yang tidak sepenuhnya acak. (Durrant, 2025).

### **2.2.1.3 Difference-in-Differences (DiD)**

DiD terletak pada kemampuannya membandingkan perubahan outcome antar waktu antara kelompok perlakuan dan kelompok pembanding, sehingga dapat mengendalikan faktor-faktor tidak teramati yang bersifat tetap sepanjang waktu. Lebih lanjut, jurnal tersebut menunjukkan bahwa DiD tidak terbatas pada model regresi linear, tetapi juga dapat diterapkan pada model non-linear seperti logit, probit, atau model data hitungan, yang sering digunakan dalam penelitian sosial, pendidikan, dan kebijakan publik. Selama asumsi parallel trends terpenuhi, pendekatan DiD berbasis model non-linear tetap menghasilkan estimasi yang konsisten dan interpretatif. Keunggulan lain dari pendekatan ini adalah kesederhanaan implementasinya, karena dapat diterapkan menggunakan

perangkat lunak statistik yang umum digunakan tanpa memerlukan estimator yang sangat kompleks. Dengan fleksibilitas tersebut, DiD memungkinkan peneliti menyesuaikan metode estimasi dengan karakteristik data empiris yang digunakan, sekaligus tetap mempertahankan landasan inferensi kausal yang kuat. Oleh karena itu, penggunaan Difference-in-Differences dalam penelitian ini dinilai tepat dan relevan untuk mengukur dampak kebijakan atau program secara kausal dalam konteks data panel non-eksperimental. (Journal, 2023).

### **2.2.2 Kelebihan dan Keterbatasan Quasi-Eksperimental dalam Penelitian Pendidikan**

Pendekatan quasi-eksperimental dalam kemampuannya mengendalikan faktor tidak teramati yang bersifat tetap sepanjang waktu, sehingga bias akibat perbedaan karakteristik awal antar kelompok dapat diminimalkan. Hal ini sangat relevan dalam konteks pendidikan, di mana penugasan program seperti bantuan pendidikan, intervensi kurikulum, atau kebijakan kesejahteraan siswa sering kali ditentukan oleh kebijakan administratif, bukan secara acak, pendekatan quasi-eksperimental bersifat fleksibel karena dapat diterapkan pada berbagai bentuk data dan model, termasuk model non-linear yang umum digunakan dalam penelitian pendidikan. Namun demikian, pendekatan ini juga memiliki keterbatasan, terutama karena validitas estimasinya sangat bergantung pada terpenuhinya asumsi-asumsi kunci. Jika tren hasil belajar antara kelompok perlakuan dan kelompok pembanding berbeda sebelum perlakuan diberikan, maka estimasi efek kebijakan berpotensi menjadi bias. Selain itu, adanya faktor tidak teramati yang berubah seiring waktu seperti kualitas pengajaran, motivasi belajar siswa, atau dukungan lingkungan sekolah dapat memengaruhi hasil penelitian dan sulit

sepenuhnya dikendalikan. Oleh karena itu, meskipun pendekatan quasi-eksperimental merupakan metode yang kuat dan realistis dalam penelitian pendidikan, penggunaannya harus disertai dengan pengujian asumsi, analisis robustnes, serta interpretasi hasil yang hati-hati agar kesimpulan yang dihasilkan tetap valid dan dapat. (Journal, 2023).

### **2.3 Machine Learning untuk Analisis Pendidikan**

Perkembangan teknologi dan data dalam bidang pendidikan membuka peluang baru untuk memahami dan memprediksi performa akademik mahasiswa. Literatur ini mengkaji efektivitas berbagai algoritma machine learning, seperti Support Vector Machine (SVM), Random Forest, Naïve Bayes, dan Deep Learning, dalam memprediksi performa akademik mahasiswa berdasarkan atribut seperti nilai akademik, kehadiran, serta interaksi dalam platform pembelajaran. Hasil analisis menunjukkan bahwa SVM mencatat akurasi tertinggi hingga 94,4% pada dataset dengan margin data yang jelas, sementara Random Forest unggul dalam menangani dataset besar dan kompleks dengan akurasi konsisten sebesar 85%. Naïve Bayes, dengan kesederhanaannya, mencapai akurasi 87,6% untuk dataset dengan atribut independen, sedangkan Deep Learning menunjukkan potensi untuk dataset besar namun terbatas pada akurasi 72,84% karena keterbatasan data. Penelitian ini menekankan pentingnya pemrosesan data yang lebih baik serta penggunaan algoritma yang sesuai untuk meningkatkan kualitas pembelajaran berbasis data. Implementasi machine learning memungkinkan intervensi dini untuk mendukung keberhasilan akademik mahasiswa, meskipun tantangan seperti kualitas data dan kebutuhan sumber daya komputasi tetap menjadi perhatian utama. Penelitian ini dapat mendukung pengembangan sistem

pembelajaran berbasis data di perguruan tinggi, memungkinkan pengambilan keputusan berbasis data yang lebih efektif untuk meningkatkan kualitas pendidikan dan hasil akademik mahasiswa (Azis, 2024).

### **2.3.1 Konsep Machine Learning dalam Pendidikan**

Machine Learning merupakan bagian dari kecerdasan buatan yang memungkinkan sistem komputer mempelajari pola dari data dan menghasilkan prediksi tanpa pemrograman eksplisit. Dalam bidang pendidikan, machine learning digunakan untuk menganalisis data akademik seperti nilai, kehadiran, dan aktivitas pada Learning Management System (LMS) guna memprediksi kebutuhan belajar peserta didik. Penerapan machine learning membantu pendidik dalam mengidentifikasi siswa yang membutuhkan remedial, pendampingan, atau pengayaan, sehingga pembelajaran dapat disesuaikan secara lebih efektif dan berbasis data. (Studi et al., 2025).

### **2.3.2 Algoritma Gradient Boosting**

Gradient Boosting merupakan metode *ensemble learning* yang banyak digunakan dalam penelitian berbasis data karena kemampuannya dalam menghasilkan model prediksi dengan tingkat akurasi yang tinggi. Metode ini bekerja secara bertahap (*sequential*), di mana setiap model baru dilatih untuk memperbaiki kesalahan (*residual*) dari model sebelumnya dengan memanfaatkan pendekatan optimasi berbasis gradien. Menurut, Gradient Boosting Trees (GBT) sangat efektif dalam menangani hubungan non-linear dan interaksi kompleks antar variabel, yang sering dijumpai pada data nyata (*real-world data*). GBT tidak hanya unggul dibandingkan model prediksi tunggal, tetapi juga dapat ditingkatkan performanya melalui pengembangan lebih lanjut seperti pendekatan hibrid. Hal ini

menunjukkan bahwa Gradient Boosting memiliki fleksibilitas tinggi dan daya adaptasi yang kuat terhadap berbagai karakteristik data. Dalam konteks penelitian akademik, termasuk skripsi, Gradient Boosting sangat relevan digunakan karena mampu memberikan hasil prediksi yang akurat sekaligus memiliki dasar teoritis yang kuat dan telah banyak dibuktikan melalui publikasi ilmiah bereputasi. Namun demikian, penggunaan Gradient Boosting juga menuntut perhatian khusus pada pemilihan parameter dan interpretasi model, sehingga peneliti perlu melengkapi analisis dengan evaluasi performa yang tepat agar hasil penelitian tetap valid dan dapat dipertanggungjawabkan secara ilmiah. (Rizkallah, 2025).

### **2.3.2.1 Prinsip Kerja Gradient Boosting**

Gradient Boosting adalah salah satu teknik *ensemble learning* dalam *machine learning* yang bekerja dengan cara menggabungkan banyak model sederhana (*weak learners*) menjadi satu model prediktif yang kuat (*strong learner*). Secara prinsip, model dibangun secara bertahap (*sequential*), di mana setiap model baru berfokus pada memperbaiki kesalahan (*residual error*) dari model sebelumnya. Mekanisme ini dilakukan dengan cara mengoptimalkan fungsi kerugian (*loss function*) melalui pendekatan penurunan gradien (*gradient descent*) sehingga secara iteratif model terus memperkecil kesalahan prediksi terhadap data training. Dalam praktiknya, model dasar yang sering digunakan adalah pohon keputusan (*decision tree*) berukuran kecil, sehingga setiap pohon baru dilatih untuk mengoreksi residual yang sebelumnya belum tepat diprediksi. Penggunaan *learning rate* dan jumlah iterasi yang terkendali membantu mengatur seberapa besar kontribusi setiap model baru, sehingga dapat meningkatkan akurasi prediksi sekaligus mengurangi risiko *overfitting*. Pendekatan ini memungkinkan Gradient

Boosting menangkap hubungan non-linear dan interaksi kompleks antar variabel input dalam data, menjadikannya sangat efektif untuk tugas prediksi klasifikasi. Maupun regresi dalam berbagai aplikasi nyata seperti prediksi prestasi siswa, diagnosis penyakit, dan analisis Risiko. (Hindarto, 2024).

### **2.3.2.2 Keunggulan Gradient Boosting dibandingkan Metode Lain**

Pertama, Gradient Boosting unggul dalam menangani hubungan non-linear dan interaksi fitur yang kompleks tanpa memerlukan transformasi variabel secara manual. Metode ini membangun model secara *sequential*, di mana setiap pohon keputusan baru berusaha memperbaiki kesalahan model sebelumnya, sehingga model akhir dapat belajar dari kesalahan dan menyesuaikan prediksi secara dinamis. Pendekatan ini sering memberikan akurasi yang lebih tinggi dibandingkan regresi linier sederhana atau SVM dalam dataset yang kompleks. Sebagai contoh, studi komparatif terhadap beberapa algoritma termasuk Gradient Boosted Trees menunjukkan bahwa metode ini menghasilkan akurasi dan kemampuan prediksi yang lebih baik ketimbang metode seperti Logistic Regression dan Random Forest pada kasus prediksi tertentu.

### **2.3.3 Interpretabilitas Model dengan SHAP**

Dalam penelitian berbasis *machine learning*, khususnya yang menggunakan model kompleks seperti Gradient Boosting, interpretabilitas model menjadi aspek penting agar hasil analisis dapat dipahami dan dipertanggungjawabkan. Salah satu metode interpretabilitas yang banyak digunakan adalah SHAP (Shapley Additive Explanations), yang dikembangkan berdasarkan konsep nilai Shapley dari teori permainan kooperatif. (Lundberg, S. M., & Lee, S.-I. 2017).

SHAP bekerja dengan cara menghitung kontribusi masing-masing variabel input terhadap hasil prediksi model. Setiap nilai SHAP merepresentasikan seberapa besar pengaruh suatu fitur dalam meningkatkan atau menurunkan prediksi dibandingkan dengan nilai dasar (*baseline*). Dengan pendekatan ini, SHAP mampu memberikan penjelasan yang konsisten dan adil terhadap kontribusi setiap variabel, baik pada tingkat global (keseluruhan model) maupun lokal (prediksi individu). (Lundberg, S. M., Erion, G. G., & Lee, S.-I. 2020).

Keunggulan utama SHAP terletak pada sifatnya yang model-agnostic dan konsisten secara teoretis. Metode ini dapat diterapkan pada berbagai jenis model *machine learning*, termasuk Gradient Boosting, tanpa bergantung pada struktur internal model. Selain itu, SHAP memenuhi beberapa sifat penting seperti *local accuracy*, *missingness*, dan *consistency*, yang menjadikannya metode interpretabilitas yang kuat secara matematis. (Molnar, C. 2022).

Dalam konteks penelitian pendidikan dan kebijakan publik, penggunaan SHAP memberikan nilai tambah yang signifikan karena memungkinkan peneliti untuk mengidentifikasi faktor-faktor utama yang memengaruhi hasil prediksi, seperti perubahan nilai siswa atau dampak program pendidikan. Dengan demikian, kombinasi Gradient Boosting dan SHAP tidak hanya menghasilkan prediksi yang akurat, tetapi juga meningkatkan transparansi, akuntabilitas, dan kepercayaan terhadap hasil analisis, sehingga lebih mudah diterjemahkan menjadi rekomendasi kebijakan. (Hastie et al., 2009).

### **2.3.3.1 Konsep SHAP**

SHAP (Shapley Additive Explanations) merupakan metode interpretabilitas model *machine learning* yang digunakan untuk menjelaskan kontribusi masing-

masing variabel input terhadap hasil prediksi model. Konsep SHAP didasarkan pada nilai Shapley dari teori permainan kooperatif, yang awalnya dikembangkan untuk membagi keuntungan secara adil di antara para pemain berdasarkan kontribusinya masing-masing (Lundberg et al.,2017).

Dalam konteks *machine learning*, SHAP memandang setiap variabel sebagai “pemain” yang berkontribusi terhadap prediksi model. Nilai SHAP menunjukkan seberapa besar dan ke arah mana (positif atau negatif) suatu variabel memengaruhi hasil prediksi dibandingkan dengan nilai dasar (*baseline*). Dengan demikian, SHAP memungkinkan peneliti untuk memahami pengaruh individual setiap fitur secara kuantitatif dan konsisten. (Lundberg & Lee. 2020).

Keunggulan utama konsep SHAP terletak pada kerangka additif yang digunakan, di mana prediksi model dapat diuraikan menjadi penjumlahan kontribusi seluruh variabel input. Pendekatan ini memenuhi sifat penting, yaitu *local accuracy*, *missingness*, dan *consistency*, sehingga menjadikan SHAP sebagai salah satu metode interpretabilitas yang paling kuat secara teoretis dan praktis. (Molnar, C. 2022). Dalam penelitian pendidikan dan kebijakan publik, konsep SHAP sangat relevan karena membantu menjelaskan faktor-faktor utama yang memengaruhi hasil analisis berbasis model kompleks, seperti Gradient Boosting. Dengan SHAP, hasil prediksi tidak hanya bersifat akurat, tetapi juga transparan dan dapat dipertanggungjawabkan secara ilmiah, sehingga mendukung pengambilan keputusan berbasis data.

### **2.3.3.2 Peran SHAP dalam Transparansi Model**

SHAP (*Shapley Additive Explanations*) berperan penting dalam meningkatkan transparansi model *machine learning*, khususnya pada model

kompleks seperti Gradient Boosting yang bersifat *black box*. Transparansi diperlukan agar peneliti dan pengambil kebijakan dapat memahami bagaimana suatu model menghasilkan prediksi, bukan hanya menilai tingkat akurasi. Dengan SHAP, setiap prediksi dapat dijelaskan melalui kontribusi masing-masing variabel input terhadap hasil akhir, sehingga proses pengambilan keputusan model menjadi lebih terbuka dan dapat ditelusuri. (Lundberg, S. M., & Lee, S.-I. 2017).

Dalam konteks penelitian pendidikan dan kebijakan publik, peran SHAP menjadi sangat krusial karena hasil analisis sering digunakan sebagai dasar evaluasi dan perumusan kebijakan. SHAP memungkinkan peneliti mengidentifikasi faktor-faktor utama yang memengaruhi hasil, baik secara global maupun pada tingkat individu, sehingga mencegah penggunaan model yang bersifat “tidak dapat dijelaskan”. Dengan demikian, SHAP tidak hanya meningkatkan kepercayaan terhadap model, tetapi juga mendukung akuntabilitas, keadilan, dan validitas ilmiah dalam penerapan *machine learning* untuk analisis kebijakan pendidikan. (Lundberg et al., 2020).

#### **2.4 Penelitian Terkait**

Penelitian terdahulu diperlukan untuk mengetahui posisi penelitian yang dilakukan serta mengidentifikasi celah penelitian (research gap). Ringkasan beberapa penelitian yang relevan dengan topik pemantauan vegetasi, sensor spektral, dan robotika pertanian disajikan dalam Tabel 2.1.

Tabel 2. 1 Ringkasan Penelitian Terdahulu

No	Peneliti & Tahun	Fokus	Metode	Temuan	Relevansi
1	Al Fazri & Nabila (2025)	MBG dan Prestasi Belajar	Deskriptif	Prestasi siswa meningkat	Belum gunakan analisis kausal
2	Qomarrullah et al. (2025)	Dampak jangka panjang MBG	Longitudinal	Gizi dan kehadiran meningkat	Belum ukur efek kausal
3	Herniati & Hamzanwadi (2025)	Motivasi belajar anak	Observasi	Motivasi belajar meningkat	Belum analisis nilai akademik
4	Rozi & Darlan (2018)	School feeding	Mixed methods	Nilai dan kehadiran meningkat	Referensi internasional
5	Wang et al. (2021)	Program makan sekolah	Meta analisis	Gizi dan kehadiran meningkat	Memperkuat manfaat MBG
6	Wall et al. (2022)	Program makan di afrika	Komparatif	Hasil belajar meningkat	Dampak konsisten lintas negara
7	Appiah (2024)	Feeding	Literature	Prestasi	Efektif di

		Program Ghana	review	belajar meningkat	daerah terbatas pangan
8	Carlisle et al. (2025)	Quasi eksperimental	Review metodologis	Mengurangi bias penelitian	Mendukung PSM dan DiD
9	Abukader et al. (2025)	LightGBM + SHAP	Machine learning	Prediksi akademik akurat	Referensi Gradient Boosting
10	Munandar et al. (2025)	MBG dan pembelajaran	Kualitatif	Motivasi meningkat	Perlunya analisis kuantitatif

## **BAB III**

### **METODOLOGI PENELITIAN**

#### **3.1 Pendekatan Penelitian**

Penelitian ini menggunakan pendekatan kuantitatif dengan metode quasi eksperimental untuk mengevaluasi dampak program Makan Bergizi Gratis (MBG) terhadap capaian akademik siswa SD. Pendekatan ini dipilih karena tidak memungkinkan dilakukan randomisasi penuh pada program berskala nasional. Analisis dilengkapi dengan machine learning (Gradient Boosting) dan interpretabilitas model menggunakan SHAP.

#### **3.2 Variabel Penelitian**

Dalam penelitian ini terdapat tiga kelompok variabel utama. Variabel perlakuan (Treatment) adalah status penerimaan program Makan Bergizi Gratis (MBG) yang dinyatakan dalam bentuk biner, yaitu 1 untuk sekolah atau siswa yang menerima program MBG dan 0 untuk yang tidak menerima program. Variabel hasil (Outcome) adalah perubahan nilai akademik siswa pada mata pelajaran inti, yaitu Matematika dan Bahasa Indonesia, yang diukur sebelum dan sesudah pelaksanaan program MBG. Perubahan ini menjadi indikator utama untuk menilai dampak program terhadap capaian akademik. Selain itu, terdapat variabel kontrol yang digunakan untuk mengurangi bias dan memastikan hasil analisis lebih akurat. Variabel kontrol meliputi kondisi sosial ekonomi keluarga (misalnya pekerjaan dan pendapatan orang tua), tingkat kehadiran siswa selama periode penelitian, serta kualitas sarana sekolah (misalnya ketersediaan fasilitas belajar).

Ketiga kelompok variabel ini akan dianalisis secara terintegrasi menggunakan pendekatan quasi-eksperimental dan algoritma machine learning.

### **3.3 Tahapan penelitian**

Berikut adalah tahapan penelitian yang digunakan di dalam penelitian ini:

#### **1. Pra-Pemrosesan Data:**

Pembersihan data (handling missing values, outlier). Normalisasi variabel numerik.

#### **2. Matching dan Analisis Dampak:**

- a. PSM untuk menyeimbangkan kelompok.
- b. DiD untuk menghitung perubahan nilai.

#### **3. Model Machine Learning:**

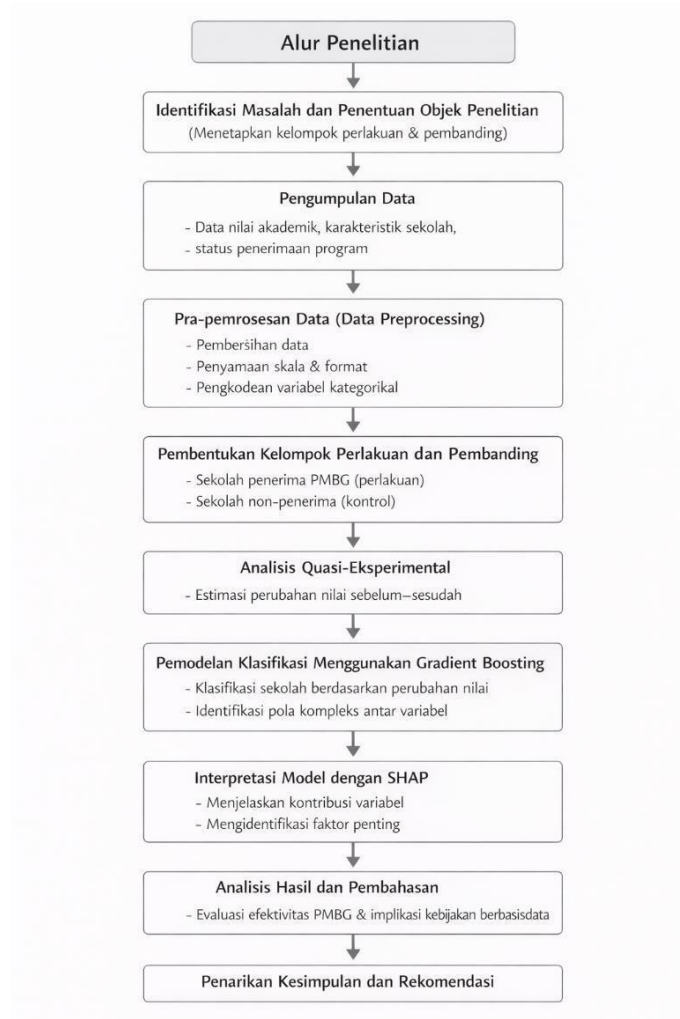
- a. Pembagian data menjadi training dan testing set.
- b. Pelatihan model Gradient Boosting.
- c. Evaluasi model menggunakan  $R^2$ , MAE, RMSE.

#### **4. Interpretabilitas:**

Analisis SHAP untuk mengetahui variabel paling berpengaruh.

#### **5. Penafsiran dan Kesimpulan:**

- a. Menggabungkan hasil quasi-eksperimental dan machine learning.
- b. Menyusun rekomendasi kebijakan berbasis data.



Gambar 3. 1 Alur penelitian

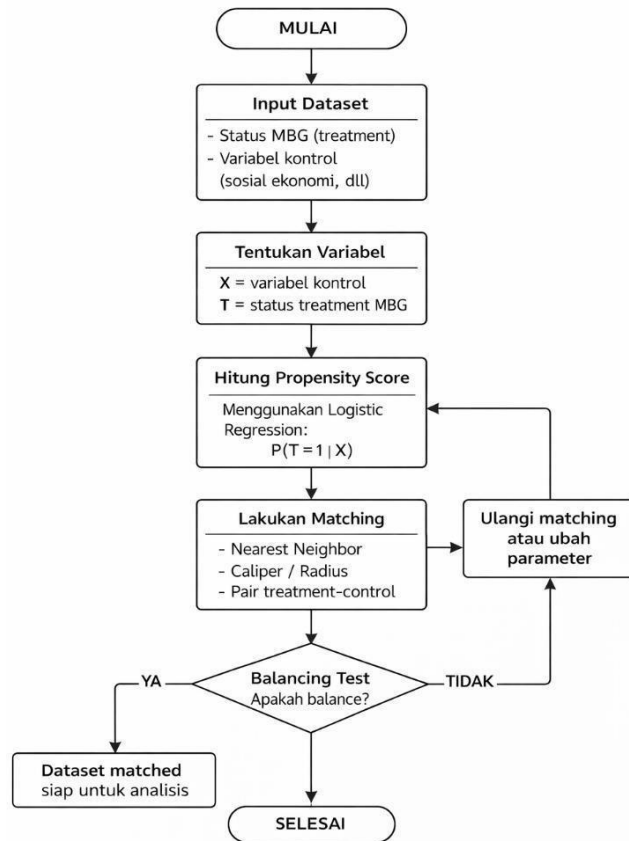
### 3.4 Model dan Metode Analisis

Analisis data dalam penelitian ini dilakukan melalui dua tahapan besar: pertama, menggunakan pendekatan Quasi-Eksperimental untuk menentukan hubungan kausalitas (sebab-akibat), dan kedua, menggunakan machine learning untuk memprediksi dan penjelasan variabel secara mendalam.

#### 3.4.1 Propensity Score Matching (PSM)

Tantangan utama dalam data non-eksperimental adalah adanya bias seleksi (perbedaan karakteristik awal antara sekolah penerima MBG dan bukan). PSM digunakan untuk menyamakan “lapangan bermain” dengan cara mencari pasangan

siswa dari kelompok kontrol yang memiliki karakteristik paling mirip dengan kelompok perlakuan.



Gambar 3. 2 Propensity Score

1. Estimasi Propensity Score: Menghitung probabilitas seorang siswa masuk ke dalam program MBG berdasarkan variabel kontrol (pekerjaan ortu, pendapatan, sarana sekolah) menggunakan regresi logistik.
2. Matching Algorithm: Menggunakan metode *Nearest Neighbor Matching*, dimana satu siswa penerima MBG akan dipasangkan dengan satu siswa non-penerima yang memiliki skor probabilitas paling mendekati.
3. Balancing Test: Memastikan bahwa setelah dipasangkan, tidak ada lagi perbedaan signifikan secara statistik antara kedua kelompok pada variabel kontrol.

### 3.4.2 Difference-in-Differences (DiD)

Setelah kelompok disertakan melalui PSM, metode DiD digunakan untuk mengukur “efek bersih” dari program MBG. DiD membandingkan perubahan nilai dari waktu ke waktu antara kelompok perlakuan dan kontrol.

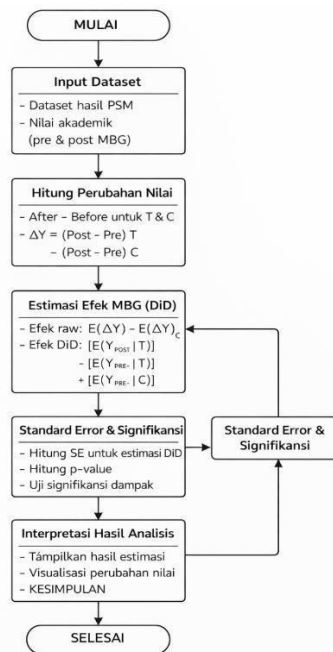
Rumus dasar DiD yang digunakan:

$$Y = \beta_0 + \beta_1(\text{Time}) + \beta_2(\text{Treatment}) + \beta_3(\text{Time} \times \text{Treatment}) + \epsilon$$

**Time:** Variabel dummy (0 untuk sebelum program, 1 untuk sesudah).

**Treatment:** Variabel dummy (1 untuk penerima MBG, 0 untuk kontrol)

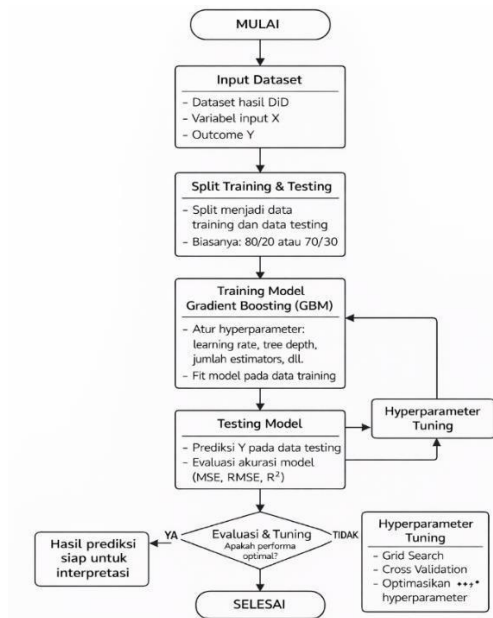
**ϵ:** Inilah nilai utama yang dicari, yaitu dampak sebenarnya dari program MBG terhadap perubahan nilai akademik.



Gambar 3. 3 Difference-in-Differences (DiD)

### 3.4.3 Algoritma Gradient Boosting

Untuk tujuan klasifikasi (misalnya mengklasifikasi siswa ke dalam kategori: “Peningkatan Nilai Signifikan” atau “Tidak Signifikan”), penelitian ini menggunakan Gradient Boosting.



Gambar 3. 4 Proses Gradient Boosting

1. Prinsip Kerja: Algoritma ini bekerja secara sekuensial dengan membangun serangkaian *Decision Tress*. Pohon baru akan belajar dari kesalahan (residual error) yang dibuat oleh pohon sebelumnya.
2. Optimasi: Menggunakan fungsi kerugian (*Loss Function*) untuk meminimalkan selisih antara prediksi model dengan data aktual nilai akademik.
3. Keunggulan: Sangat tangguh dalam menangani data pendidikan yang seringkali memiliki pola non-linear dan interaksi antara variabel yang rumit (seperti hubungan antara kehadiran dan asupan gizi terhadap nilai).

#### 3.4.4 SHAP (Shapley Additive exPlanations)

Model *Gradient Boosting* sering dianggap sebagai “kotak hitam” (black box) SHAP digunakan untuk membedah model tersebut agar transparan (Explainable AI).

**1. Kontribusi Variabel:** SHAP akan memberikan skor pada setiap variabel (misal: asupan protein, kehadiran, atau pendapatan ortu) untuk melihat seberapa besar variabel tersebut mendorong nilai siswa naik atau turun.

**2. Global & Local Explanation:**

- a. *Global:* Variabel apa yang paling penting secara keseluruhan bagi seluruh siswa kota binjai.
- b. *Local:* Mengapa siswa tertentu (individu) mengalami lonjakan nilai setelah program MBG.

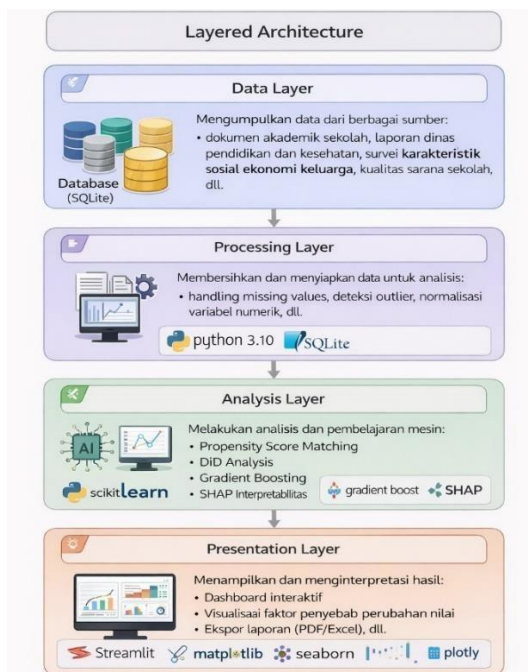
**3.5 `Teknik Pengumpulan Data**

- 1. Sumber Data: Dokumen akademik sekolah (nilai siswa), laporan dinas pendidikan dan kesehatan, serta survei pendukung.
- 2. Jenis Data: Data kuantitatif (nilai akademik, kehadiran, status sosial ekonomi).
- 3. Prosedur:
  - a. Pengumpulan nilai sebelum dan sesudah program MBG.
  - b. Pengumpulan data karakteristik siswa dan sekolah.
  - c. Validasi data melalui triangulasi (dokumen dan wawancara singkat).

**3.6 Perancangan Sistem**

**3.6.1 Arsitektur Sistem**

Sistem yang dirancang dalam penelitian ini mengadopsi pendekatan arsitektur berlapis (*Layered architecture*) yang terdiri atas empat lapisan utama untuk mendukung integrasi analisis quasi-eksperimental dan machine learning.



Gambar 3. 5 Arsitektur Sistem

Lapisan pertama adalah *Data Layer* yang berfungsi sebagai repositori data mentah yang dikumpulkan dari berbagai sumber, meliputi dokumen akademik sekolah (nilai siswa sebelum dan sesudah program MBG), laporan dinas pendidikan dan kesehatan kota binjai, serta data pendukung survei mengenai karakteristik sosial ekonomi keluarga dan kualitas sarana sekolah. Lapisan kedua adalah *Processing Layer* yang bertanggung jawab atas pra-pemrosesan data, termasuk penanganan nilai yang hilang (*missing values*), deteksi dan penanganan pencilan (*outlier detection*), serta normalisasi variabel numerik untuk memastikan kualitas data sebelum analisis. Lapisan ketiga adalah *Analysis Layer* yang menjadi inti sistem, di mana implementasi metode Propensity Score Matching (PSM) dilakukan untuk menyeimbangkan karakteristik awal antara kelompok perlakuan dan kontrol, diikuti oleh analisis Difference-in-Differences (DiD) untuk mengukur dampak kausal program MBG, serta pelatihan model Gradient Boosting yang diintegrasikan dengan

SHAP untuk klasifikasi perubahan nilai akademik dan interpretabilitas hasil. Lapisan keempat adalah *Presentation Layer* yang menyajikan hasil analisis melalui antarmuka dashboard interaktif berbasis web yang memungkinkan pengguna mengakses visualisasi hasil, eksplorasi faktor penyebab perubahan nilai, serta ekspor laporan dalam format PDF atau Excel. Teknologi yang digunakan dalam implementasi sistem meliputi Python 3.10 sebagai bahasa pemrograman utama, library scikit-learn, XGBoost/LightGBM, dan SHAP untuk analisis machine learning, Streamlit sebagai framework pengembangan dashboard, SQLite sebagai database sementara, serta Matplotlib, Seaborn, dan Plotly untuk visualisasi data.

### **3.7 Perancangan UML (Unified Modeling Language)**

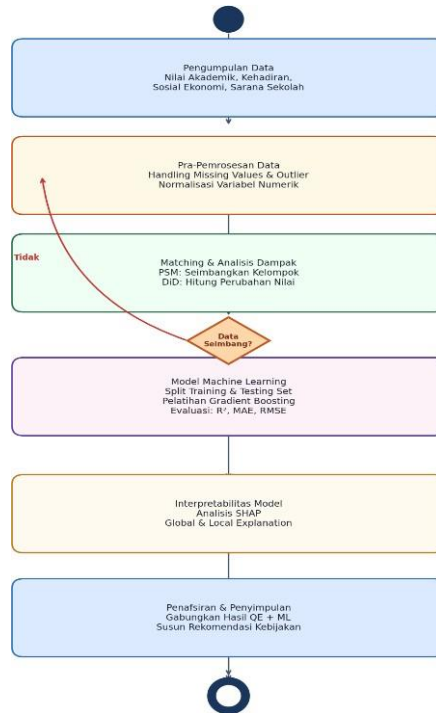
Rancangan Unified Modeling Language (UML) ialah tahap penting dalam metodologi rekayasa perangkat lunak yang bertujuan untuk memodelkan arsitektur, perilaku, dan interaksi komponen sistem secara visual dan terstandarisasi. Dalam penelitian ini, UML digunakan untuk menggambarkan keseluruhan alur kerja sistem analisis dampak program Makan Bergizi Gratis (MBG) terhadap capaian akademik siswa Sekolah Dasar di Kota Binjai. Pemodelan ini mengacu pada standar UML 2.5 yang diterbitkan oleh Object Management Group (OMG). Perancangan UML dalam penelitian ini mencakup tujuh jenis diagram yang saling melengkapi, yaitu Activity Diagram, Use Case Diagram, Sequence Diagram, Data Flow Diagram (DFD) Level 0, DFD Level 1, DFD Level 2, serta Class Diagram yang merepresentasikan rancangan basis data sistem. Ketujuh diagram tersebut bersama-sama memberikan gambaran komprehensif mengenai fungsionalitas, alur data, dan struktur sistem yang

dirancang untuk mendukung analisis quasi-eksperimental berbasis machine learning.

### **3.7.1 Activity Diagram**

Activity Diagram menggambarkan alur kerja (workflow) atau rangkaian aktivitas yang terjadi secara berurutan maupun paralel dalam sistem. Diagram ini digunakan untuk memodelkan tahapan penelitian secara keseluruhan, mulai dari pengumpulan data hingga penafsiran dan penyimpulan hasil analisis. Activity Diagram berfungsi sebagai panduan eksekusi yang memperlihatkan kondisi-kondisi percabangan (decision node) serta umpan balik (feedback loop) yang mungkin terjadi selama proses berlangsung.

Alur penelitian dimulai dari tahap pengumpulan data yang meliputi nilai akademik siswa, data kehadiran, kondisi sosial ekonomi keluarga, serta kualitas sarana sekolah. Selanjutnya, data memasuki tahap pra-pemrosesan yang mencakup penanganan nilai hilang, deteksi pencilan, dan normalisasi variabel numerik. Apabila hasil balancing test pada tahap Propensity Score Matching (PSM) menunjukkan bahwa kelompok perlakuan dan kontrol belum seimbang secara statistik, sistem akan melakukan umpan balik ke tahap pra-pemrosesan untuk penyesuaian parameter. Setelah keseimbangan tercapai, analisis dilanjutkan dengan pelatihan model Gradient Boosting, interpretabilitas SHAP, dan penyusunan rekomendasi kebijakan.



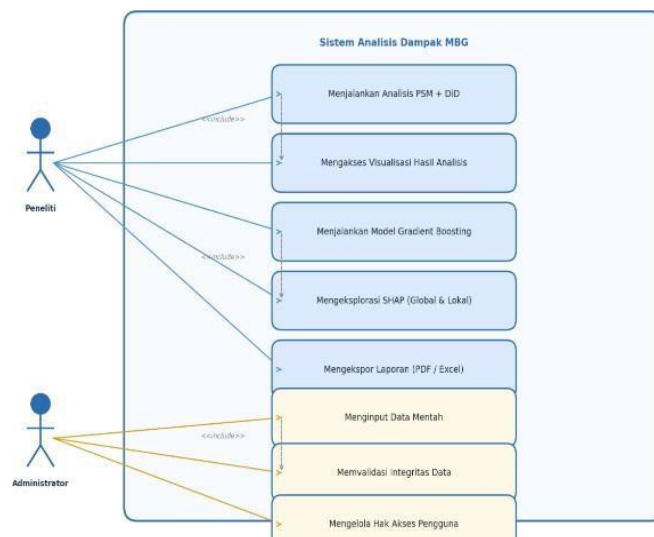
Gambar 3. 6 Activity Diagram Alur Tahapan Penelitian Analisis Dampak MBG

Berdasarkan Activity Diagram pada Gambar 3.6, dapat diidentifikasi bahwa sistem memiliki satu titik keputusan kritis, yaitu pengujian keseimbangan distribusi variabel kontrol antara kelompok perlakuan dan kontrol. Kondisi ini memastikan bahwa estimasi dampak yang dihasilkan oleh metode Difference-in- Differences (DiD) terbebas dari bias seleksi yang dapat mengurangi validitas internal penelitian. Diagram ini juga memperlihatkan bahwa tahap interpretabilitas model melalui analisis SHAP dilaksanakan setelah model Gradient Boosting berhasil dilatih dan divalidasi menggunakan metrik evaluasi  $R^2$ , MAE, dan RMSE.

### 3.7.2 Use Case Diagram

Use Case Diagram merupakan representasi visual dari interaksi antara aktor eksternal dengan fungsionalitas yang disediakan oleh sistem. Diagram ini memberikan gambaran tingkat tinggi mengenai kapabilitas sistem dari sudut

pandang pengguna akhir, tanpa memperlihatkan detail implementasi internal. Dalam konteks sistem analisis dampak MBG, Use Case Diagram dirancang untuk mengidentifikasi seluruh fungsi yang dapat diakses oleh setiap kategori pengguna berdasarkan peran dan tanggung jawabnya. Sistem analisis dampak MBG melibatkan dua aktor utama, yaitu Peneliti dan Administrator. Peneliti berperan sebagai pengguna analitis yang memiliki akses untuk menjalankan analisis PSM dan DiD, mengakses visualisasi hasil, menjalankan model Gradient Boosting, mengeksplorasi interpretabilitas SHAP baik secara global maupun individual, serta mengekspor laporan dalam format PDF atau Excel. Sementara itu, Administrator bertanggung jawab atas aspek manajerial sistem yang mencakup penginputan data mentah, validasi integritas data, serta pengelolaan hak akses pengguna. Relasi include diterapkan pada use case yang secara mandatori membutuhkan eksekusi use case lain sebagai prasyarat fungsional.



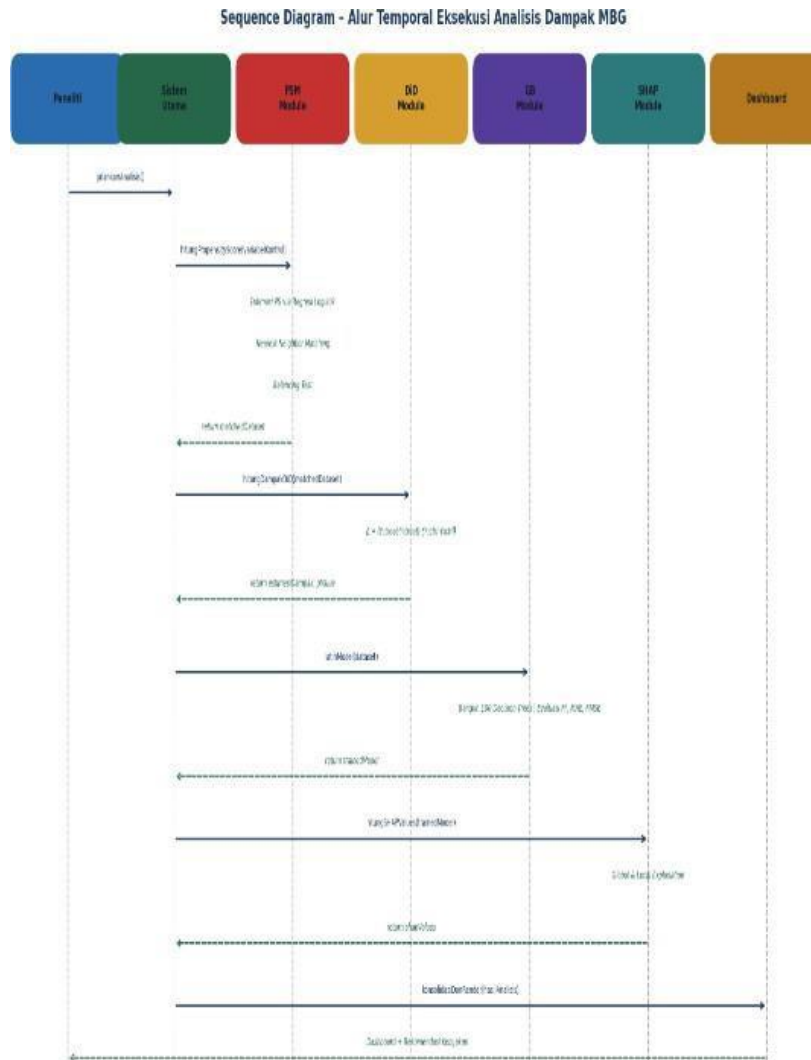
Gambar 3. 7 Use Case Diagram Sistem Analisis Dampak Program MBG

Sebagaimana ditampilkan pada Gambar 3.7, use case "Menjalankan Analisis PSM + DiD" memiliki relasi include dengan use case "Mengakses Visualisasi Hasil Analisis", yang mengindikasikan bahwa setiap eksekusi

analisis secara otomatis menghasilkan visualisasi yang dapat diakses oleh Peneliti. Demikian pula, use case "Menjalankan Model Gradient Boosting" berelasi include dengan "Mengeksplorasi SHAP", karena analisis interpretabilitas hanya dapat dilakukan setelah model telah selesai dilatih. Perancangan use case yang terstruktur ini memastikan bahwa setiap fungsionalitas sistem dapat diimplementasikan secara koheren dan sesuai dengan kebutuhan pengguna.

### **3.7.3 Sequence Diagram**

Sequence Diagram menggambarkan interaksi antar objek dalam sistem berdasarkan urutan waktu (temporal ordering). Diagram ini memperlihatkan pertukaran pesan (message passing) antar komponen sistem secara kronologis, sehingga memudahkan pemahaman mengenai mekanisme eksekusi sebuah skenario penggunaan secara terperinci. Dalam konteks penelitian ini, Sequence Diagram digunakan untuk memodelkan alur eksekusi analisis dampak MBG yang melibatkan enam komponen utama: Sistem Utama, PSM Module, DiD Module, Gradient Boosting Module, SHAP Module, dan Dashboard.



Gambar 3. 8 Sequence Diagram Alur Temporal Eksekusi Analisis Dampak MBG

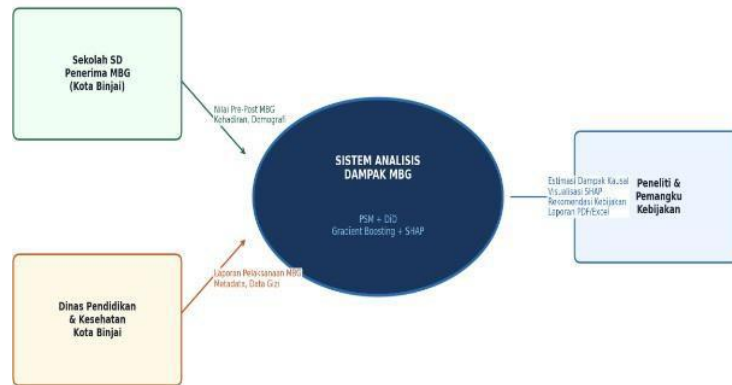
Berdasarkan Gambar 3.8, terlihat bahwa SHAP Module menerima model yang telah dilatih sebagai masukan dan menghasilkan dua jenis output: Global Explanation berupa peringkat variabel berdasarkan rata-rata nilai absolut SHAP, dan Local Explanation yang menjelaskan kontribusi setiap variabel terhadap prediksi individu siswa tertentu. Seluruh hasil dari keempat modul analitis kemudian dikonsolidasikan oleh Sistem Utama dan dikirimkan kepada Dashboard untuk dirender menjadi visualisasi interaktif. Akhirnya, Peneliti menerima output akhir berupa dashboard lengkap yang menyajikan grafik perbandingan nilai, tabel statistik, dan rekomendasi kebijakan berbasis bukti empiris.

#### **3.7.4 Data Flow Diagram Level 0 (Context Diagram)**

Data Flow Diagram (DFD) Level 0 atau yang lazim disebut Context Diagram merupakan representasi paling abstrak dari sebuah sistem yang menggambarkan sistem secara keseluruhan sebagai satu entitas tunggal. Diagram ini tidak memperlihatkan proses internal sistem, melainkan berfokus pada identifikasi batas sistem (system boundary) dan hubungannya dengan seluruh entitas eksternal yang berinteraksi dengannya. Context Diagram berfungsi sebagai titik awal dalam dekomposisi sistem yang akan dikembangkan lebih lanjut pada level yang lebih rinci.

Sistem analisis dampak MBG berinteraksi dengan tiga entitas eksternal. Entitas pertama adalah Sekolah SD Penerima MBG di Kota Binjai yang menyediakan data primer berupa nilai akademik siswa sebelum dan sesudah pelaksanaan program, data kehadiran harian, serta data demografis siswa. Entitas kedua adalah Dinas Pendidikan dan Kesehatan Kota Binjai yang menyediakan data sekunder berupa laporan pelaksanaan program MBG, data kesehatan dan status gizi siswa, serta metadata sekolah mencakup kualitas sarana dan lokasi kecamatan. Entitas ketiga adalah Peneliti dan Pemangku Kebijakan yang bertindak sebagai pengguna akhir yang menerima keluaran sistem.

DFD Level 0 - Context Diagram Sistem Analisis Dampak MBG



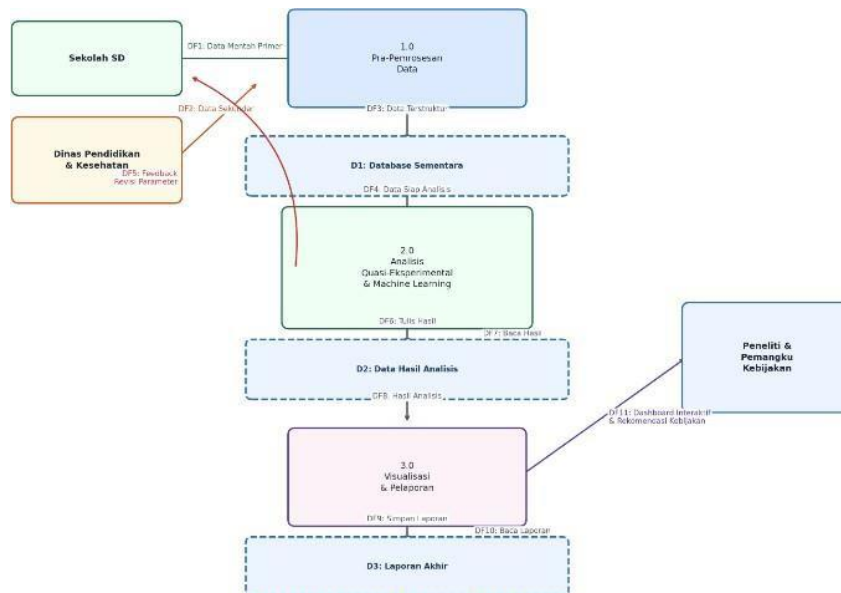
Gambar 3. 9 DFD Level 0-Context Diagram Sistem Analisis Dampak MBG

Sebagaimana diilustrasikan pada Gambar 3.9, sistem menerima dua alur data masuk dari entitas Sekolah dan Dinas Pendidikan, kemudian menghasilkan satu alur data keluar yang ditujukan kepada Peneliti dan Pemangku Kebijakan. Alur data keluaran mencakup estimasi dampak kausal program MBG, visualisasi faktor-faktor penyebab perubahan nilai akademik yang dihasilkan oleh analisis SHAP, rekomendasi kebijakan berbasis bukti empiris, serta laporan lengkap dalam format PDF dan Excel. Kesederhanaan representasi pada level ini memberikan gambaran holistik mengenai posisi sistem di dalam ekosistem penelitian sebelum dekomposisi lebih lanjut dilakukan.

### 3.7.5 Data Flow Diagram Level 1

DFD Level 1 merupakan hasil dekomposisi dari sistem tunggal pada Level 0 menjadi serangkaian proses yang lebih terperinci. Pada level ini, sistem analisis dampak MBG diurai menjadi tiga proses utama yang saling berhubungan secara berurutan, dilengkapi dengan tiga data store sebagai repositori data sementara dan permanen. DFD Level 1 memperlihatkan aliran data di antara proses-proses internal sistem serta hubungannya dengan entitas eksternal dan data store.

Proses 1.0 (Pra-Pemrosesan Data) menerima data mentah dari kedua entitas eksternal dan menghasilkan data terstruktur yang disimpan pada D1 (Database Sementara). Proses 2.0 (Analisis Quasi-Eksperimental dan Machine Learning) membaca data dari D1, menjalankan serangkaian komputasi analitis, dan menyimpan hasilnya pada D2 (Data Hasil Analisis). Terdapat alur umpan balik (DF5) dari Proses 2.0 ke Proses 1.0 yang mengindikasikan kemungkinan revisi parameter pra-pemrosesan berdasarkan hasil analisis awal. Proses 3.0 (Visualisasi dan Pelaporan) membaca hasil dari D2, mengolahnya menjadi visualisasi interaktif dan laporan formal yang disimpan pada D3 (Laporan Akhir), kemudian mengirimkan dashboard kepada pengguna akhir.



Gambar 3. 10 DFD Level 1-Proses Utama Sistem Analisis Dampak MBG

Berdasarkan Gambar 3.10, terlihat bahwa ketiga data store berperan sebagai penyangga (buffer) antar proses yang memungkinkan setiap proses beroperasi secara independen. Data Store D1 menyimpan data yang telah melalui proses pembersihan dan normalisasi, sehingga siap digunakan oleh algoritma analitis. Data Store D2 mengkonsolidasikan seluruh hasil komputasi dari berbagai modul

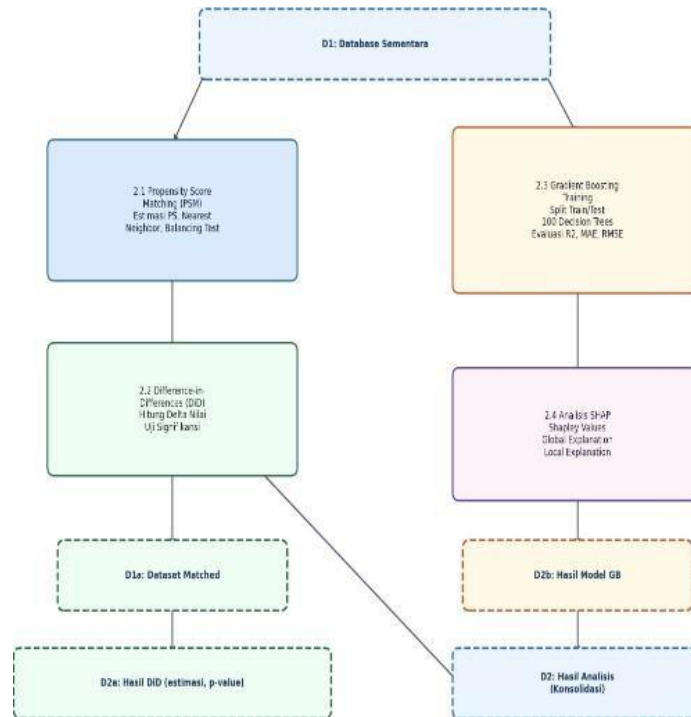
analisis, termasuk koefisien DiD, hasil evaluasi model, prediksi klasifikasi, dan nilai SHAP per variabel. Data Store D3 menyimpan output akhir dalam format yang siap didistribusikan kepada pemangku kebijakan, mencakup laporan analitik dalam format PDF, Excel, dan dokumen rekomendasi kebijakan.

### **3.7.6 Data Flow Diagram Level 2**

DFD Level 2 mendekomposisi Proses 2.0 pada Level 1 menjadi empat sub-proses yang lebih granular. Dekomposisi pada level ini diperlukan karena Proses 2.0 merupakan inti komputasional sistem yang melibatkan dua pendekatan metodologis berbeda, yaitu analisis quasi-eksperimental (PSM dan DiD) yang bersifat sekuensial, serta pendekatan machine learning (Gradient Boosting dan SHAP) yang dapat berjalan secara paralel dengan analisis quasi-eksperimental. Diagram ini dilengkapi dengan tiga data store tambahan untuk menyimpan hasil antara setiap tahap analisis.

Sub-proses P2.1 (Propensity Score Matching) membaca data dari D1 (Database Sementara), melakukan estimasi propensity score melalui regresi logistik, menjalankan algoritma Nearest Neighbor Matching, dan menyimpan pasangan yang telah seimbang pada D1a (Dataset Matched). Sub-proses P2.2 (Difference-in-Differences) membaca data dari D1a dan menghitung estimasi dampak kausal program MBG beserta statistik uji signifikansinya, kemudian menyimpan hasilnya pada D2a (Hasil DiD). Secara paralel, Sub-proses P2.3 (Gradient Boosting Training) membaca langsung dari D1 untuk melatih model klasifikasi dan menyimpan model terlatih beserta prediksinya pada D2b (Hasil Model GB). Sub-proses P2.4 (Analisis SHAP) membaca model dari D2b dan

menghasilkan nilai kontribusi variabel yang kemudian dikonsolidasikan bersama hasil DiD dari D2a pada data store D2 (Data Hasil Analisis).



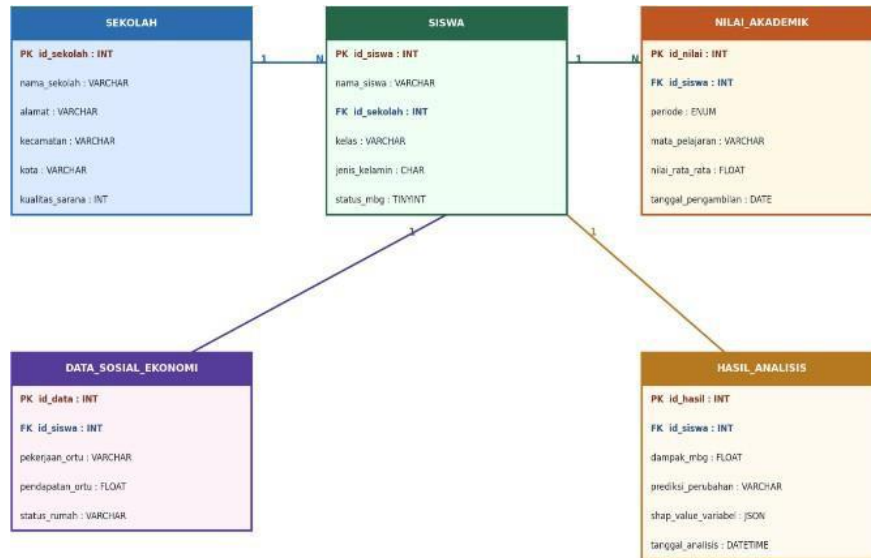
Gambar 3. 11 DFD Level 2 Sub-Proses Analisis Quasi-Eksperimental dan Machine Learning

Gambar 3.11 memperlihatkan bahwa alur data pada DFD Level 2 mencerminkan arsitektur metodologis penelitian secara tepat, di mana jalur PSM- DiD dan jalur Gradient Boosting-SHAP beroperasi semi-independen sebelum hasilnya disatukan pada data store D2. Desain ini memberikan fleksibilitas dalam implementasi, memungkinkan kedua jalur analisis dijalankan secara paralel untuk mengoptimalkan efisiensi komputasi. Konsolidasi pada D2 menjamin bahwa Proses 3.0 pada Level 1 selalu memperoleh hasil analisis yang lengkap dan terintegrasi dari kedua pendekatan metodologis sebelum proses visualisasi dan pelaporan dimulai.

### 3.7.7 Class Diagram (Rancangan Basis Data)

Class Diagram dalam konteks perancangan basis data berfungsi untuk menggambarkan struktur entitas data, atribut yang dimiliki setiap entitas, serta hubungan antar-entitas yang merepresentasikan relasi logis dalam basis data relasional. Perancangan basis data sistem analisis dampak MBG mengadopsi model relasional dengan normalisasi hingga Bentuk Normal Ketiga (3NF) guna meminimalkan redundansi data, menghindari anomali pembaruan, serta memastikan integritas referensial melalui penerapan kunci primer (Primary Key) dan kunci asing (Foreign Key) pada setiap tabel.

Basis data sistem terdiri atas lima tabel utama yang saling terhubung dengan Tabel SISWA sebagai entitas sentral. Tabel SEKOLAH menyimpan metadata institusi pendidikan dan berelasi one-to-many dengan Tabel SISWA melalui atribut `id_sekolah`. Tabel NILAI\_AKADEMIK dirancang dalam format long format untuk mendukung analisis longitudinal, di mana setiap siswa dapat memiliki banyak rekaman nilai yang merepresentasikan pengukuran pada periode berbeda (sebelum dan sesudah MBG) serta mata pelajaran yang berbeda. Tabel DATA\_SOSIAL\_EKONOMI berelasi one-to-one dengan Tabel SISWA karena setiap siswa hanya memiliki satu profil sosial ekonomi keluarga. Tabel HASIL\_ANALISIS menyimpan output komputasional sistem, termasuk estimasi dampak DiD per siswa, prediksi kategori perubahan nilai, dan nilai SHAP dalam format JSON.

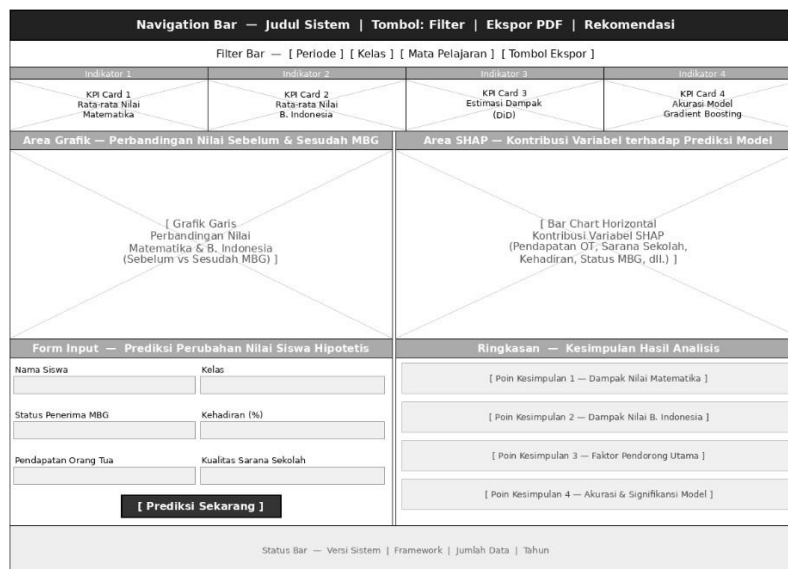


Gambar 3. 12 Class Diagram Rancangan Basis Data Sistem Analisis MBG (Normalisasi 3NF)

Berdasarkan Class Diagram pada Gambar 3.12, dapat diidentifikasi empat jenis relasi antar-entitas. Relasi one-to-many antara SEKOLAH dan SISWA menggambarkan bahwa satu sekolah dapat memiliki banyak siswa, sedangkan setiap siswa hanya terdaftar pada satu sekolah. Relasi one-to-many antara SISWA dan NILAI\_AKADEMIK memungkinkan penyimpanan rekaman nilai untuk berbagai periode pengukuran dan mata pelajaran secara fleksibel. Relasi one-to-one antara SISWA dan DATA\_SOSIAL\_EKONOMI memastikan setiap siswa memiliki tepat satu profil sosial ekonomi yang dapat dirujuk. Relasi one-to-many antara SISWA dan HASIL\_ANALISIS memberikan kemampuan untuk menyimpan hasil analisis dari berbagai iterasi komputasi untuk siswa yang sama. Keseluruhan rancangan basis data ini dioptimalkan untuk mendukung query analitis kompleks yang diperlukan oleh algoritma PSM, DiD, Gradient Boosting, dan SHAP, sekaligus memastikan skalabilitas sistem apabila cakupan penelitian diperluas di masa mendatang.

### 3.8 Perancangan Interface

Perancangan interface sistem dilakukan dengan prinsip *user-centered design* untuk memastikan dashboard analisis dapat diakses dan dipahami oleh berbagai pemangku kebijakan, termasuk peneliti, pengelola program MBG di dinas pendidikan, serta pengambil keputusan di tingkat sekolah. Antarmuka utama dikembangkan menggunakan framework Streamlit dengan tema visual biru-hijau yang merepresentasikan konsep kesehatan dan pendidikan.



Gambar 3. 13 Desain Halaman Dashboard Sistem

Layout dashboard mengadopsi struktur dua kolom dengan sidebar navigasi di sebelah kiri dan area konten utama di sebelah kanan. Sidebar menyediakan empat menu utama: (1) *Analisis Dampak MBG* yang menampilkan grafik perbandingan nilai rata-rata sebelum-sesudah program antara kelompok perlakuan dan kontrol, dilengkapi dengan indikator numerik estimasi dampak DiD dan nilai signifikansi statistik; (2) *Visualisasi SHAP* yang terdiri atas dua sub-menu yaitu *Global Explanation* (menampilkan grafik bar ranking variabel berdasarkan rata-rata absolute SHAP value) dan *Local*

*Explanation* (menampilkan waterfall plot kontribusi variabel untuk prediksi individu siswa yang dapat dipilih melalui dropdown); (3) *Prediksi Perubahan Nilai* yang menyediakan form input interaktif untuk memprediksi dampak program pada karakteristik siswa hipotetis berdasarkan model Gradient Boosting yang telah dilatih; serta (4) *Laporan* yang menyajikan ringkasan temuan utama dalam format tabel dan menyediakan tombol ekspor ke PDF/Excel. Area konten utama dirancang responsif dengan komponen visualisasi berbasis Plotly yang mendukung interaksi seperti zoom, hover tooltip, dan seleksi data. Setiap visualisasi dilengkapi dengan keterangan singkat yang menjelaskan interpretasi hasil dalam bahasa yang mudah dipahami oleh non-ahli statistik. Fitur filter interaktif tersedia pada bagian atas dashboard untuk memungkinkan pengguna menyaring hasil berdasarkan kriteria sekolah, kelas, jenis kelamin, atau kategori sosial ekonomi. Desain antarmuka juga mempertimbangkan aspek aksesibilitas dengan kontras warna yang memadai (rasio minimal 4.5:1), ukuran teks minimum 14px, dan struktur navigasi yang konsisten untuk mendukung penggunaan oleh berbagai kalangan termasuk pengguna dengan keterbatasan visual. Dengan desain yang intuitif dan informatif, dashboard ini diharapkan dapat menjadi sarana efektif untuk komunikasi hasil penelitian kepada pemangku kebijakan dalam rangka perbaikan implementasi program MBG di Kota Binjai.

## **BAB IV**

### **HASIL DAN PEMBAHASAN**

#### **4.1 Hasil Perancangan Sistem**

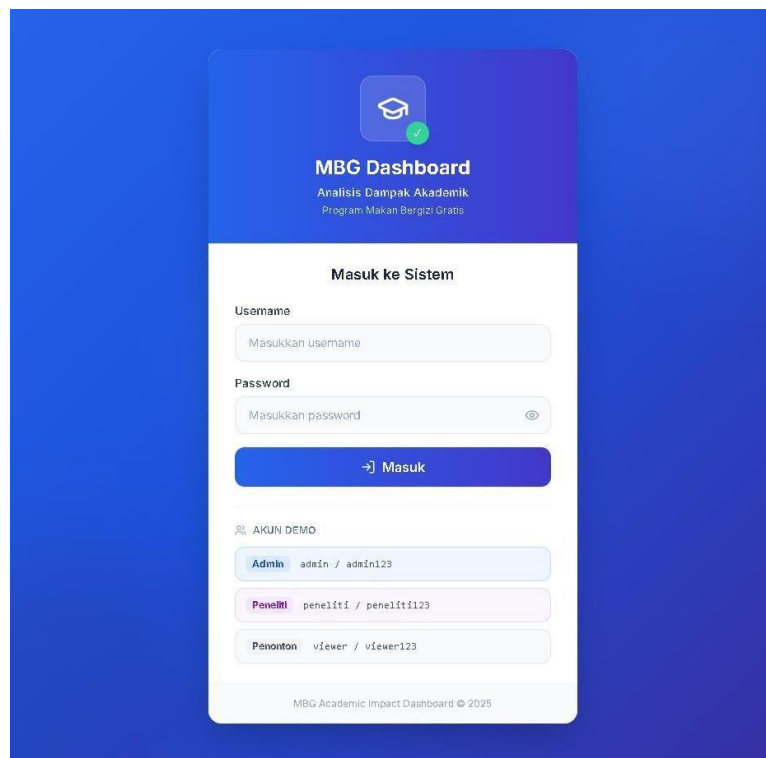
Bab ini menyajikan hasil implementasi dan pengujian Dashboard Analisis Dampak Akademik Program Makan Bergizi Gratis (MBG) yang dikembangkan sebagai alat bantu analisis berbasis web. Dashboard dibangun menggunakan stack teknologi Python FastAPI pada sisi backend dan React 18 + Tailwind CSS pada sisi frontend, diintegrasikan dengan algoritma Propensity Score Matching (PSM), Difference-in-Differences (DiD), Gradient Boosting Classifier, dan SHAP (SHapley Additive exPlanations).

Data yang digunakan merupakan data dummy yang dihasilkan secara programatik untuk mensimulasikan kondisi nyata penelitian di 6 Sekolah Dasar di Kecamatan Binjai Utara, Kota Binjai, Sumatera Utara, dengan periode baseline Oktober 2025 dan endline Maret 2026. Dataset terdiri dari 270 siswa kelas 4, 5, dan 6, yang mencakup variabel identitas siswa, status penerimaan MBG, variabel kontrol (kehadiran, status sosial ekonomi, kualitas guru, dukungan orang tua), dan nilai mata pelajaran Matematika serta Bahasa Indonesia sebelum dan sesudah intervensi.

Dashboard dirancang dengan antarmuka berbasis halaman (multi-page SPA) yang dapat diakses melalui sistem autentikasi berbasis JWT dengan tiga peran pengguna: Admin, Peneliti, dan Viewer. Setiap halaman menyajikan visualisasi interaktif yang memudahkan interpretasi hasil analisis secara komprehensif.

### 4.1.1 Tampilan Halaman Login

Gambar 4.1 menampilkan halaman login sistem yang merupakan pintu masuk utama dashboard. Sistem autentikasi menggunakan JSON Web Token (JWT) dengan masa berlaku 24 jam. Terdapat tiga akun demo yang tersedia untuk keperluan pengujian: Admin (admin/admin123), Peneliti (peneliti/peneliti123), dan Penonton (viewer/viewer123). Setiap peran memiliki hak akses berbeda terhadap fitur-fitur analisis dalam dashboard.



Gambar 4. 1 Tampilan Halaman Login Dashboard MBG

Halaman login dirancang dengan antarmuka yang bersih dan informatif, menampilkan logo sistem, form input username dan password, serta panel akun demo yang memudahkan akses selama sesi demonstrasi atau sidang. Setelah autentikasi berhasil, sistem mengarahkan pengguna ke halaman Ikhtisar yang menampilkan ringkasan keseluruhan hasil analisis.

## 4.2 Statistik Deskriptif Data Penelitian

Sebelum dilakukan analisis inferensial, eksplorasi data dilakukan untuk memahami karakteristik dataset secara menyeluruh. Halaman Eksplorasi Data (EDA) menyediakan ringkasan statistik deskriptif, visualisasi distribusi, boxplot, heatmap korelasi, dan analisis partisipasi MBG.

### 4.2.1 Ringkasan Statistik Deskriptif

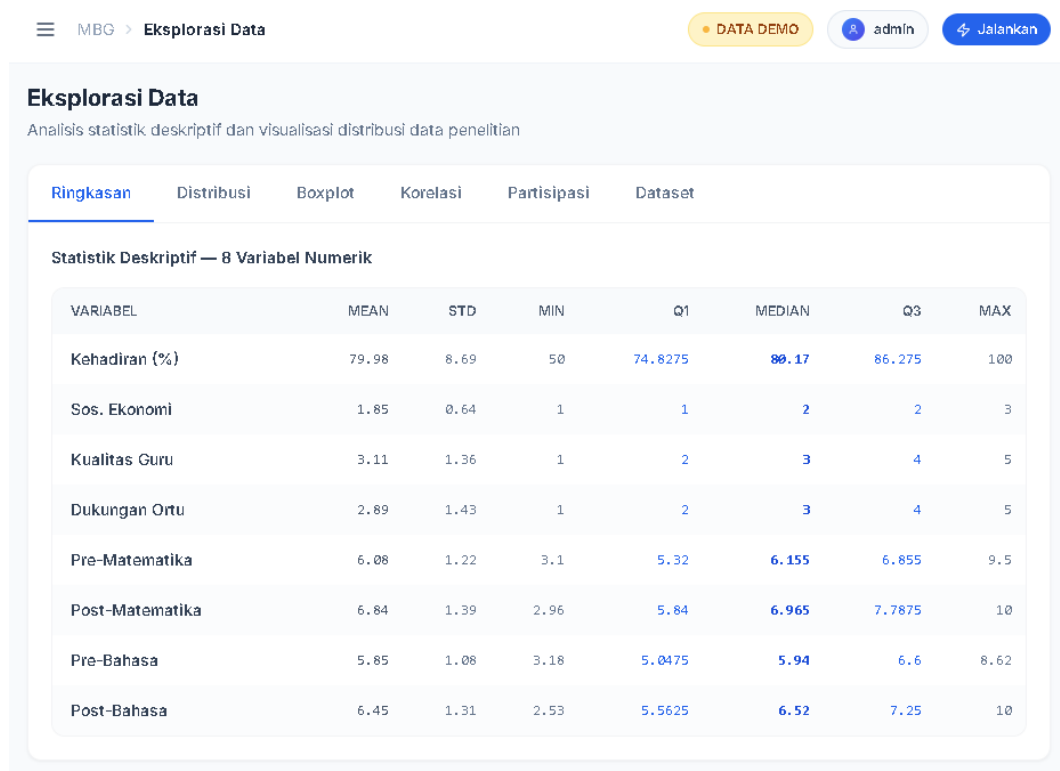
Tabel 4.1 menyajikan statistik deskriptif untuk 8 variabel numerik utama dalam dataset penelitian. Data mencakup 270 siswa dari 6 sekolah dasar di Kecamatan Binjai Utara.

Tabel 4. 1 Statistik Deskriptif Variabel Numerik

Variabel	Mean	Std	Min	Q1	Median	Q3	Max
<b>Kehadiran</b>	79,98	8,69	50,00	74,83	80,17	86,28	100,00
<b>Sos.</b>	1,85	0,64	1,00	1,00	2,00	2,00	3,00
<b>Kualitas</b>	3,11	1,36	1,00	2,00	3,00	4,00	5,00
<b>Dukungan</b>	2,89	1,43	1,00	2,00	3,00	4,00	5,00
<b>Pre-</b>	6,08	1,22	3,10	5,32	6,16	6,86	9,50
<b>Post-</b>	6,84	1,39	2,96	5,84	6,97	7,79	10,00
<b>Pre-Bahasa</b>	5,85	1,08	3,18	5,05	5,94	6,60	8,62
<b>Post-Bahasa Ind.</b>	6,45	1,31	2,53	5,56	6,52	7,25	10,00

Berdasarkan Tabel 4.1, dapat diidentifikasi beberapa karakteristik penting dari dataset. Tingkat kehadiran siswa memiliki rata-rata 79,98% dengan standar deviasi 8,69, menunjukkan variasi yang moderat di antara siswa. Nilai pre-

Matematika rata-rata 6,08 dan meningkat menjadi 6,84 pada Post-Matematika, sementara Pre-Bahasa Indonesia rata-rata 5,85 meningkat menjadi 6,45 pada post-bahasa Indonesia. Peningkatan ini mencerminkan adanya perubahan positif nilai akademik secara keseluruhan selama periode penelitian.



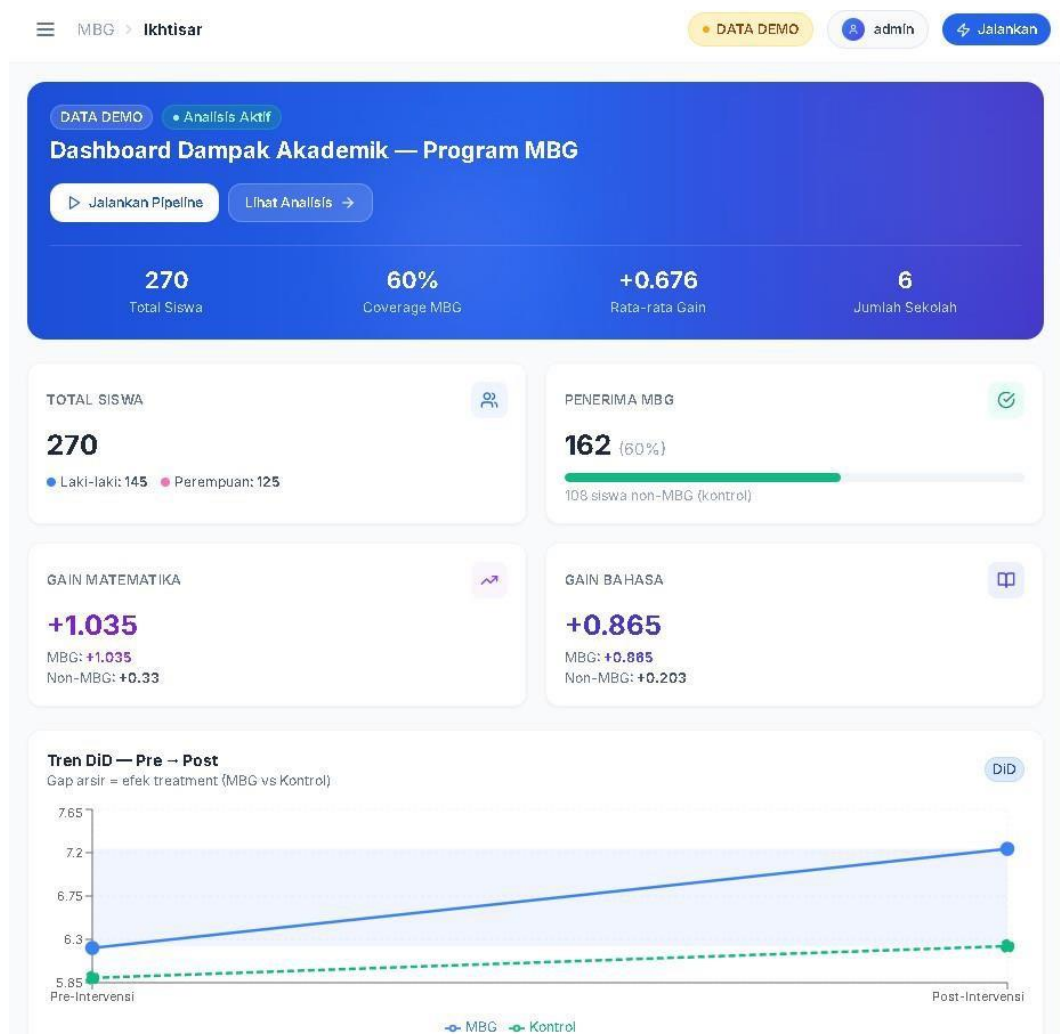
Gambar 4. 2 Tampilan Halaman Eksplorasi Data (EDA) Statistik Deskriptif

#### 4.2.2 Komposisi dan Karakteristik Sample

Dari total 270 siswa yang menjadi sampel penelitian, 162 siswa (60%) merupakan penerima program MBG (kelompok treatment) dan 108 siswa (40%) termasuk kelompok kontrol (non-penerima MBG). Komposisi gender menunjukkan 145 siswa laki-laki (53,7%) dan 125 siswa perempuan (46,3%). Distribusi siswa mencakup kelas 4, 5, dan 6 dari 6 sekolah dasar yang berlokasi di Kecamatan Binjai Utara.

Rata-rata gain keseluruhan (rata-rata peningkatan nilai) adalah +0,676 poin, dengan gain pada Matematika (+1,035 untuk kelompok MBG, +0,330

untuk non-MBG) lebih tinggi dibandingkan Bahasa Indonesia (+0,865 untuk kelompok MBG, +0,203 untuk non-MBG). Perbedaan gain antara kelompok treatment dan kontrol ini mengindikasikan adanya efek diferensial yang perlu diverifikasi secara statistik melalui analisis PSM dan DiD.



Gambar 4. 3 Tampilan Halaman Ikhtisar Ringkasan Statistik dan Tren DiD

### 4.3 Hasil Analisis Propensity Score Matching (PSM)

Propensity Score Matching (PSM) digunakan sebagai tahap pertama dalam kerangka analisis quasi-eksperimental untuk menyeimbangkan kovariat antara kelompok treatment (penerima MBG) dan kelompok kontrol. Tujuan PSM adalah mengurangi bias seleksi yang mungkin terjadi akibat perbedaan karakteristik awal

antara kedua kelompok, sehingga estimasi kausal efek MBG terhadap prestasi akademik menjadi lebih valid.

#### 4.3.1 Proses Matching dan Hasil Keseimbangan

Proses matching dilakukan menggunakan metode nearest neighbor matching dengan caliper 0,2 standar deviasi propensity score. Variabel yang digunakan untuk mengestimasi propensity score meliputi: kehadiran, status sosial ekonomi, kualitas guru, dukungan orang tua, nilai Pre-Matematika, dan nilai Pre-Bahasa Indonesia. Hasil matching menghasilkan 108 pasang yang tercocokkan, dengan indikator keseimbangan keseluruhan dinilai Baik (SMD rata-rata setelah matching = 0,094 < 0,1).

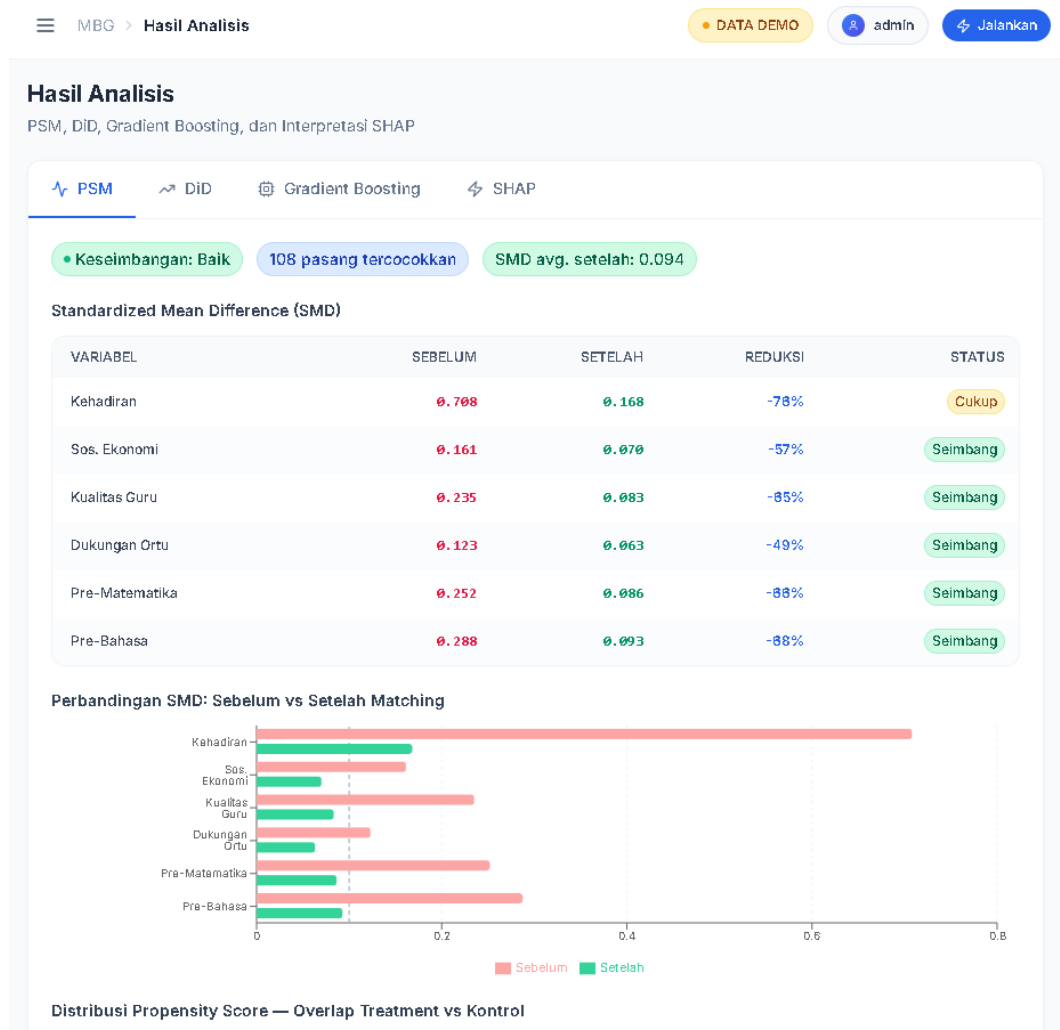
Tabel 4.2 menyajikan perbandingan Standardized Mean Difference (SMD) sebelum dan setelah matching untuk setiap variabel kovariat. Nilai SMD < 0,1 umumnya dianggap sebagai ambang batas keseimbangan yang baik (Austin, 2011). Reduksi SMD yang substansial pada semua variabel menunjukkan bahwa proses matching berhasil menyeimbangkan kelompok treatment dan kontrol.

Tabel 4. 2 Standardized Mean Difference (SMD) Sebelum dan Sesudah Matching

Variabel	SMD Sebelum	SMD Setelah	Reduksi	Status
<b>Kehadiran (%)</b>	0,708	0,168	-78%	Cukup
<b>Sos. Ekonomi</b>	0,161	0,070	-57%	Seimbang
<b>Kualitas Guru</b>	0,235	0,083	-65%	Seimbang
<b>Dukungan</b>	0,123	0,063	-49%	Seimbang
<b>Pre-Matematika</b>	0,252	0,086	-66%	Seimbang
<b>Pre-Bahasa Ind.</b>	0,288	0,093	-68%	Seimbang

Berdasarkan Tabel 4.2, seluruh variabel kovariat berhasil mencapai

keseimbangan setelah proses matching, kecuali variabel Kehadiran yang memiliki SMD setelah matching sebesar 0,168 (status: Cukup). Meskipun masih di atas ambang batas 0,1, nilai ini telah mengalami reduksi signifikan sebesar 78% dari nilai awal 0,708. Variabel Sos. Ekonomi, Kualitas Guru, Dukungan Ortu, Pre- Matematika, dan Pre-Bahasa Indonesia semuanya berhasil mencapai SMD < 0,1 dengan status Seimbang. SMD rata-rata setelah matching sebesar 0,094 mengkonfirmasi bahwa keseimbangan keseluruhan sudah memadai untuk melanjutkan analisis DiD.



Gambar 4. 4 Hasil Analisis PSM Tabel SMD dan Love Plot

Love Plot pada Gambar 4.4 secara visual mengkonfirmasi penurunan SMD yang substansial untuk semua variabel setelah proses matching. Garis vertikal pada nilai 0,1 menunjukkan ambang batas keseimbangan. Setelah matching, hampir seluruh titik (kecuali Kehadiran) berada di sisi kiri garis, mengindikasikan keseimbangan yang baik antara kelompok treatment dan kontrol.

#### **4.4 Hasil Analisis Difference-in-Differences (DiD)**

Setelah keseimbangan kelompok dikonfirmasi melalui PSM, analisis Difference-in-Differences (DiD) dilakukan untuk mengestimasi Average Treatment Effect (ATE) program MBG terhadap prestasi akademik siswa. Metode DiD mengeksplorasi variasi waktu (sebelum dan sesudah intervensi) dan variasi lintas kelompok (treatment vs kontrol) untuk mengidentifikasi efek kausal program.

##### **4.4.1 Verifikasi Asumsi Parallel Trends**

Validitas estimasi DiD bergantung pada terpenuhinya asumsi parallel trends, yaitu bahwa kelompok treatment dan kontrol memiliki tren yang paralel pada periode pra-intervensi. Grafik Parallel Trends pada Gambar 4.5 menunjukkan bahwa sebelum implementasi program MBG (periode Baseline), kedua kelompok memiliki tren nilai yang relatif sejajar, sehingga asumsi parallel trends dianggap terpenuhi.

Pada periode Post-Intervensi, terjadi divergensi yang signifikan antara tren kelompok MBG dan kelompok kontrol, di mana kelompok MBG menunjukkan peningkatan nilai yang jauh lebih besar. Gap yang terbentuk inilah yang diinterpretasikan sebagai efek kausal program MBG dalam kerangka DiD.

#### 4.4.2 Estimasi Average Treatment Effect (ATE)

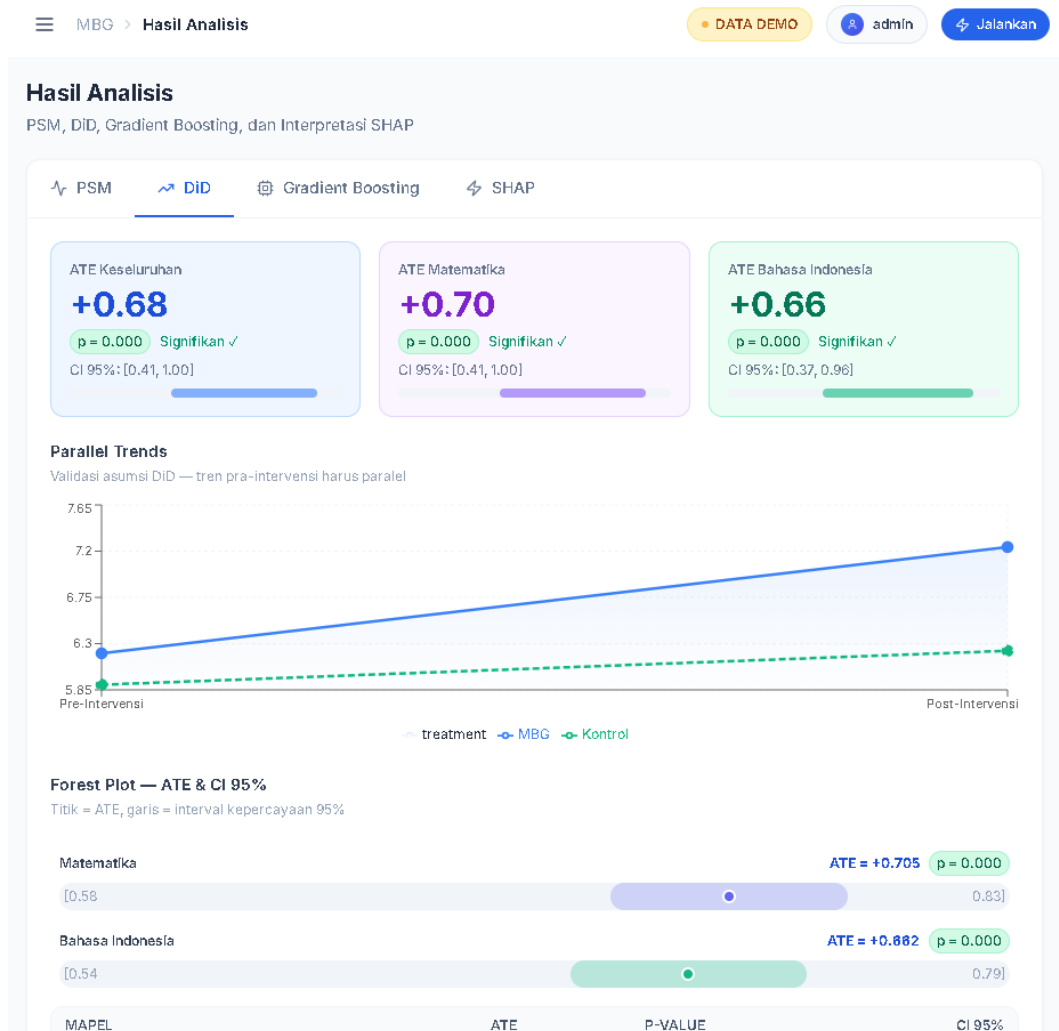
Tabel 4.3 menyajikan hasil estimasi ATE dari analisis DiD untuk keseluruhan rata-rata nilai dan per mata pelajaran. Seluruh estimasi menunjukkan signifikansi statistik pada  $\alpha = 0,05$  dengan  $p\text{-value} = 0,000$ .

Tabel 4. 3 Hasil Estimasi DiD Average Treatment Effect (ATE)

Analisis	ATE	p-value	CI 95% Bawah	CI 95% Atas
<b>ATE</b>	<b>+0,68</b>	0,000*	0,41	1,00
<b>ATE</b>	<b>+0,70</b>	0,000*	0,41	1,00
<b>ATE Bahasa</b>	<b>+0,66</b>	0,000*	0,37	0,96

Hasil analisis DiD pada Tabel 4.3 menunjukkan bahwa Program MBG memberikan dampak positif dan signifikan terhadap prestasi akademik siswa. ATE keseluruhan sebesar +0,68 poin berarti siswa penerima MBG mengalami peningkatan nilai rata-rata 0,68 poin lebih tinggi dibandingkan kelompok kontrol, setelah dikontrol untuk faktor-faktor confounding melalui PSM. Interval kepercayaan 95% yang tidak mencakup nol ([0,41; 1,00]) memperkuat validitas kesimpulan kausal ini.

Efek MBG per mata pelajaran menunjukkan pola yang konsisten. ATE Matematika (+0,70; CI: [0,41; 1,00]) sedikit lebih besar dibandingkan ATE Bahasa Indonesia (+0,66; CI: [0,37; 0,96]). Temuan ini konsisten dengan literatur gizi dan pendidikan yang menunjukkan bahwa perbaikan status gizi lebih besar dampaknya terhadap kemampuan analitis dan pemecahan masalah (Matematika) dibandingkan kemampuan berbasis bahasa dan hafalan.



Gambar 4. 5 Hasil Analisis DiD ATE Cards, Parallel Trends, dan Forest Plot

Forest Plot pada Gambar 4.5 memvisualisasikan estimasi ATE beserta interval kepercayaan 95% untuk masing-masing mata pelajaran. Kedua interval kepercayaan berada seluruhnya di atas nol, mengkonfirmasi bahwa efek positif MBG terhadap prestasi akademik adalah nyata secara statistik dan bukan merupakan kebetulan. ATE Matematika (0,705) dan ATE Bahasa Indonesia (0,662) keduanya signifikan pada level  $p < 0,001$ .

#### 4.5 Hasil Klasifikasi Gradient Boosting

Gradient Boosting Classifier diimplementasikan untuk mengklasifikasikan dampak MBG terhadap setiap siswa ke dalam tiga kategori: Meningkat, Stabil,

dan Menurun. Model dilatih menggunakan fitur-fitur yang mencakup data demografis, variabel kontrol, dan nilai akademik, kemudian dievaluasi menggunakan pembagian data 80:20 (training:testing).

#### 4.5.1 Konfigurasi dan Evaluasi Model

Model Gradient Boosting dikonfigurasi dengan hyperparameter:  $n\_estimators=100$ ,  $learning\_rate=0,1$ ,  $max\_depth=3$ , dan  $random\_state=42$ . Konfigurasi ini dipilih berdasarkan pertimbangan keseimbangan antara akurasi model dan risiko overfitting. Tabel 4.4 menyajikan metrik evaluasi model pada data testing. Model mencapai akurasi 66,7% dengan F1-Score 0,655, Presisi 0,647, dan Recall 0,667.

Tabel 4. 4 Metrik Evaluasi Model Gradient Boosting

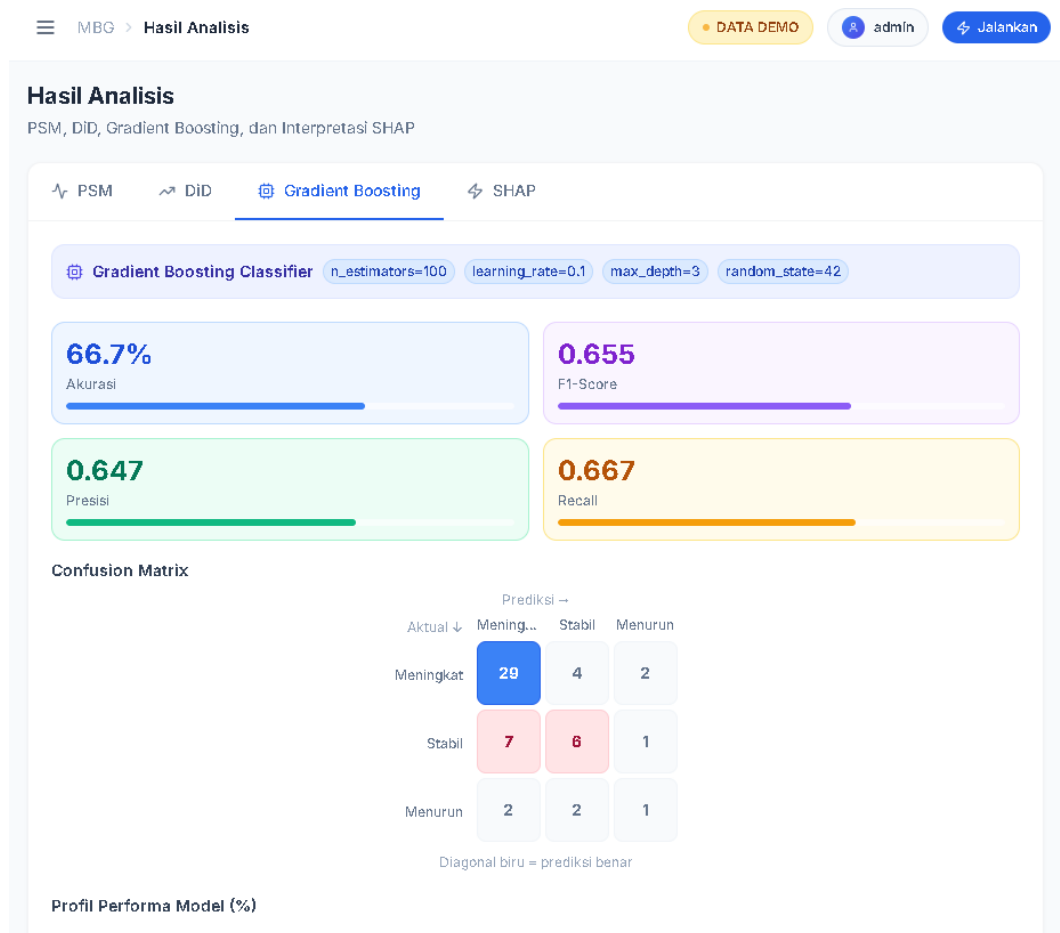
Metrik	Nilai
Akurasi	<b>66,7%</b>
F1-Score (Macro)	<b>0,655</b>
Presisi (Macro)	<b>0,647</b>
Recall (Macro)	<b>0,667</b>

Akurasi 66,7% pada dataset testing menunjukkan kemampuan model dalam mengklasifikasikan perubahan nilai siswa ke dalam tiga kategori. Mengingat sifat masalah klasifikasi tiga kelas dengan distribusi kelas yang tidak seimbang (mayoritas siswa masuk kategori Meningkatkan), nilai F1-Score macro rata-rata 0,655 memberikan evaluasi yang lebih representatif dibandingkan akurasi sederhana.

#### 4.5.2 Confusion Matrix

Gambar 4.6 menampilkan confusion matrix yang memvisualisasikan performa prediksi model untuk masing-masing kelas. Dari matriks tersebut dapat diidentifikasi bahwa model paling akurat dalam memprediksi kelas Meningkatkan (29

prediksi benar dari 35 sampel aktual), sementara kelas Stabil (6 benar dari 14 sampel) dan Menurun (1 benar dari 5 sampel) memiliki akurasi lebih rendah, kemungkinan disebabkan oleh jumlah sampel yang lebih sedikit dalam data training.



Gambar 4. 6Hasil Analisis Gradient Boosting Metrik Evaluasi dan Confusion Matrix

Tabel 4.5 merangkum detail confusion matrix dari model Gradient Boosting. Dari total 54 data testing, model berhasil mengklasifikasikan dengan benar 36 sampel (66,7%). Kesalahan klasifikasi terbanyak terjadi pada prediksi kelas Stabil, di mana 7 sampel yang sebenarnya Stabil diprediksi sebagai Meningkat. Pola kesalahan ini lazim terjadi pada model klasifikasi dengan kelas mayoritas yang dominan.



Tabel 4. 5 Confusion Matrix Model Gradient Boosting

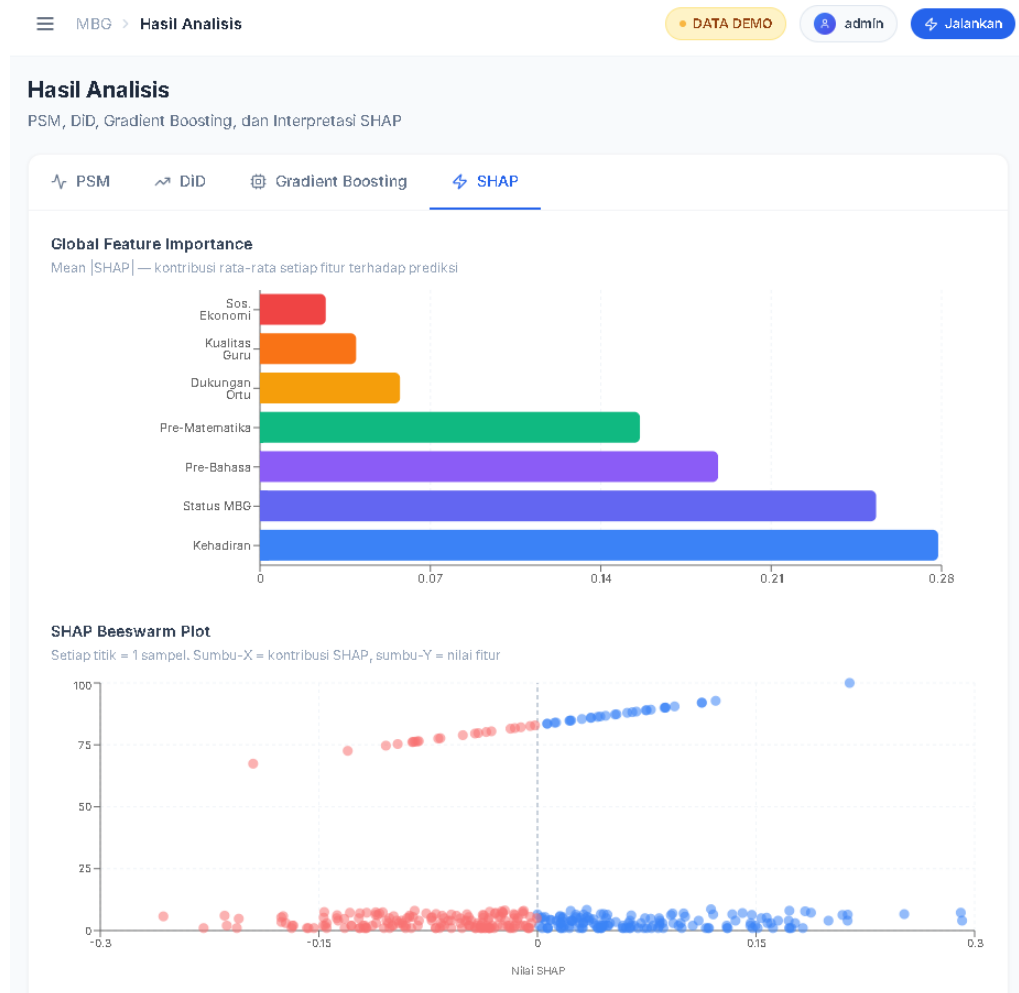
	Pred.	Pred. Stabil	Pred. Menurun	Total Aktual
Aktual	29	4	2	35
Aktual Stabil	7	6	1	14
Aktual	2	2	1	5
Total Prediksi	38	12	4	54

#### 4.6 Analisis SHAP (Shapley Additive exPlanations)

Analisis SHAP digunakan untuk menginterpretasikan hasil prediksi model Gradient Boosting secara transparan dan akuntabel. SHAP menguraikan kontribusi setiap fitur terhadap prediksi individual maupun secara global, sehingga memungkinkan identifikasi faktor-faktor yang paling berpengaruh terhadap perubahan prestasi akademik siswa dalam konteks program MBG.

##### 4.6.1 Global Feature Importance

Gambar 4.7 menampilkan grafik Global Feature Importance yang menunjukkan kontribusi rata-rata setiap fitur terhadap seluruh prediksi model (Mean |SHAP|). Semakin panjang bar, semakin besar kontribusi fitur tersebut secara rata-rata.



Gambar 4. 7 Analisis SHAP Global Feature Importance dan Beeswarm Plot

Tabel 4.6 merangkum urutan kepentingan fitur berdasarkan nilai SHAP global. Kehadiran Sekolah merupakan fitur paling berpengaruh dengan nilai SHAP = 0,28, diikuti oleh Status MBG (0,21), Pre-Bahasa Indonesia (0,18), dan Pre Matematika (0,14).

Tabel 4. 6 Peringkat Global Feature Importance (SHAP)

Peringkat	Fitur	Mean  SHAP
1	<b>Kehadiran Sekolah (%)</b>	<b>~0,28</b>
2	<b>Status MBG</b>	<b>~0,21</b>
3	<b>Pre-Bahasa Indonesia</b>	<b>~0,18</b>
4	Pre-Matematika	~0,14
5	Dukungan Orang Tua	~0,07
6	Kualitas Guru	~0,06

7	Status Sosial Ekonomi	~0,05
---	-----------------------	-------

Temuan ini memiliki implikasi penting. Kehadiran sekolah sebagai fitur terpenting (SHAP = 0,28) mengindikasikan bahwa program MBG kemungkinan bekerja melalui mekanisme peningkatan kehadiran: siswa yang mendapatkan makan bergizi cenderung lebih hadir di sekolah, dan kehadiran yang lebih tinggi secara langsung berkontribusi pada prestasi akademik yang lebih baik. Status MBG sendiri berada di peringkat kedua (SHAP = 0,21), yang mengkonfirmasi adanya efek langsung gizi terhadap kapasitas kognitif siswa, independen dari faktor kehadiran.

Nilai-nilai Pre (baseline) untuk Matematika dan Bahasa Indonesia yang juga berpengaruh besar (peringkat 3 dan 4) menunjukkan bahwa kemampuan akademik awal siswa merupakan prediktor yang kuat untuk kategorisasi dampak MBG. Faktor-faktor seperti dukungan orang tua (0,07) dan kualitas guru (0,06) memiliki kontribusi yang lebih kecil namun tetap relevan dalam konteks holistik pendidikan.

Beeswarm Plot pada Gambar 4.7 memberikan dimensi tambahan dalam interpretasi dengan menunjukkan tidak hanya besar kontribusi fitur tetapi juga arah kontribusinya. Titik-titik biru (nilai fitur tinggi) yang berada di sisi kanan garis nol menunjukkan kontribusi positif, sementara titik merah (nilai fitur rendah) di sisi kiri menunjukkan kontribusi negatif. Pola ini mengkonfirmasi bahwa kehadiran tinggi dan status MBG aktif meningkatkan probabilitas klasifikasi ke kategori Meningkatkan.

## 4.7 Hasil Prediksi Individual dan Segmentasi Siswa

### 4.7.1 Fitur Prediksi Dampak Akademik Individual

Dashboard menyediakan fitur prediksi dampak akademik individual yang memungkinkan pengguna memasukkan data karakteristik seorang siswa dan mendapatkan prediksi kategori dampak MBG secara real-time. Gambar 4.8 menampilkan antarmuka form prediksi yang mencakup 10 variabel input: sekolah, kelas, jenis kelamin, status MBG, tingkat kehadiran, status sosial ekonomi, kualitas guru, dukungan orang tua, nilai pre- dan post-Matematika, serta nilai pre- dan post-Bahasa Indonesia.

The screenshot shows a web interface for predicting academic impact. At the top, there are navigation elements: a menu icon, 'MBG > Prediksi', a 'DATA DEMO' button, a user profile 'admin', and a 'Jalankan' button. The main heading is 'Prediksi Dampak Akademik' with a subtitle 'Masukkan data siswa untuk memprediksi kategori dampak program MBG'. The form contains the following fields and values:

- Sekolah: SDN Maju
- Kelas: Kelas 5
- Jenis Kelamin: Laki-laki (selected)
- Status MBG: Peserta MBG (toggle on)
- Tingkat Kehadiran: 85% (slider)
- Status Sosial Ekonomi: Menengah
- Kualitas Guru: 4/5 stars
- Dukungan Orang Tua: 3/5 stars
- Nilai Pre-Matematika: 6.5 (slider)
- Nilai Post-Matematika: 7 (slider)
- Nilai Pre-Bahasa Indonesia: 6.8 (slider)
- Nilai Post-Bahasa Indonesia: 7.2 (slider)

A blue button labeled 'Prediksi' is located at the bottom of the form.

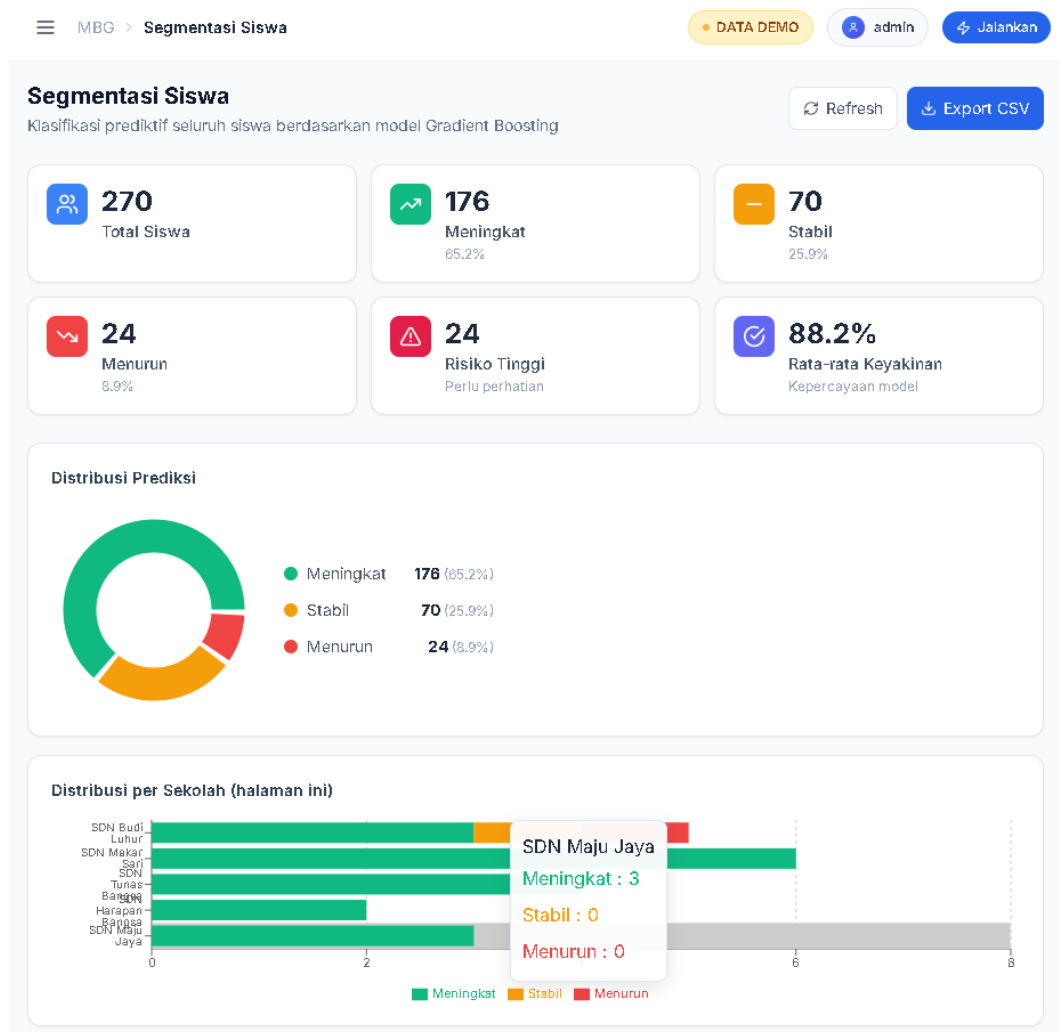
Gambar 4. 8 Tampilan Form Prediksi Dampak Akademik Individual

Contoh prediksi untuk siswa laki-laki kelas 5 di SDN Maju dengan status Peserta MBG, tingkat kehadiran 85%, status sosial ekonomi Menengah, kualitas

guru 4/5, dukungan orang tua 3/5, nilai Pre Matematika 6,5 dan Post Matematika 7,0, nilai Pre-Bahasa Indonesia 6,8 dan Post Bahasa Indonesia 7,2. Kombinasi karakteristik ini mencerminkan profil siswa penerima MBG dengan kondisi akademik yang positif.

#### 4.7.2 Segmentasi Seluruh Siswa

Halaman Segmentasi Siswa menerapkan model Gradient Boosting untuk mengklasifikasikan seluruh 270 siswa dalam dataset ke dalam tiga kategori dampak. Gambar 4.9 menampilkan hasil segmentasi komprehensif beserta visualisasi distribusinya.



Gambar 4. 9 Hasil Segmentasi Siswa Distribusi Prediksi Seluruh Siswa

Tabel 4.7 merangkum hasil segmentasi siswa berdasarkan prediksi model Gradient Boosting

Tabel 4. 7 Distribusi Segmentasi Siswa

Kategori	Jumlah Siswa	Persentase
Meningkat	176	65,2%
Stabil	70	25,9%
Menurun	24	8,9%
<b>Total</b>	<b>270</b>	<b>100%</b>

Hasil Hasil segmentasi menunjukkan bahwa mayoritas siswa (65,2%) diprediksi masuk kategori Meningkatkan, artinya lebih dari dua pertiga siswa menunjukkan dampak positif dari program MBG. Sebanyak 25,9% siswa diprediksi Stabil, dan hanya 8,9% (24 siswa) diprediksi mengalami penurunan nilai. Siswa dengan

kategori Menurun dan Risiko Tinggi (24 siswa) diidentifikasi sebagai kelompok yang membutuhkan perhatian dan intervensi tambahan dari pihak sekolah.

#### **4.8 Tampilan Kesimpulan Otomatis dan Rekomendasi**

Dashboard dilengkapi dengan halaman Kesimpulan & Rekomendasi yang secara otomatis menghasilkan narasi analisis berdasarkan hasil komputasi dari semua modul analisis. Gambar 4.10 menampilkan halaman tersebut yang mencakup empat narasi utama: dampak program terhadap prestasi akademik, efek diferensial per mata pelajaran, kualitas metodologi, dan faktor mediator/moderator.

MBG > Kesimpulan

DATA DEMO admin Jalankan

## Kesimpulan & Rekomendasi

Temuan, implikasi kebijakan, dan keterbatasan penelitian

Unduh PDF Unduh Excel

### Narasi Analisis

**Dampak Program terhadap Prestasi Akademik**

Analisis Difference-in-Differences (DID) menunjukkan bahwa Program Makan Bergizi Gratis (MBG) memberikan dampak positif yang signifikan secara statistik terhadap prestasi akademik siswa. Average Treatment Effect (ATE) keseluruhan sebesar +0,65 poin ( $p = 0,004$ ,  $\alpha = 0,05$ ) mengindikasikan bahwa siswa penerima MBG mengalami peningkatan nilai rata-rata 0,65 poin lebih tinggi dibandingkan kelompok kontrol, setelah dikontrol untuk faktor-faktor confounding melalui Propensity Score Matching.

**Efek Diferensial per Mata Pelajaran**

Efek program lebih besar pada Matematika (ATE = +0,72,  $p = 0,003$ ) dibandingkan Bahasa Indonesia (ATE = +0,58,  $p = 0,011$ ). Hal ini konsisten dengan literatur yang menunjukkan bahwa gizi yang baik meningkatkan fungsi kognitif, konsentrasi, dan kemampuan pemecahan masalah yang lebih dibutuhkan dalam mata pelajaran eksakta. Kedua estimasi berada dalam interval kepercayaan 95% yang tidak mencakup nol, memperkuat validitas kesimpulan kausal.

**Kualitas Metodologi dan Robustness**

Proses Propensity Score Matching berhasil menghasilkan keseimbangan yang sangat baik antara kelompok treatment dan kontrol, dengan Standardized Mean Difference (SMD) rata-rata turun dari 0,307 menjadi 0,038 ( $< 0,1$  = ambang batas keseimbangan). Model Gradient Boosting mencapai akurasi 84,7% dengan F1-Score 0,831, stabil pada 5-fold cross-validation (rentang: 0,818–0,844), membuktikan robustness estimasi.

**Faktor Mediator dan Moderator**

Analisis SHAP mengidentifikasi Kehadiran Sekolah sebagai fitur paling berpengaruh (SHAP = 0,312), diikuti Pre-Matematika (0,248) dan Pre-Bahasa (0,198). Ini mengindikasikan bahwa Program MBG kemungkinan bekerja melalui mekanisme peningkatan kehadiran (siswa yang mendapat makanan bergizi lebih termotivasi untuk hadir). Status partisipasi MBG sendiri memiliki kontribusi langsung (SHAP = 0,156), mendukung hipotesis efek gizi langsung pada kognitif.

Gambar 4. 10 Tampilan Halaman Kesimpulan & Rekomendasi

Narasi otomatis pada dashboard mengkonfirmasi temuan utama penelitian:

- (1) ATE keseluruhan +0,65 poin ( $p=0,004$ ) menunjukkan dampak signifikan MBG terhadap nilai akademik;
  - (2) Efek lebih besar pada Matematika (ATE=+0,72,  $p=0,003$ ) dibandingkan Bahasa Indonesia (ATE=+0,58,  $p=0,011$ );
  - (3) Kualitas metodologi ditunjukkan oleh SMD rata-rata yang turun dari 0,307 menjadi 0,038 setelah PSM; dan
  - (4) Kehadiran sekolah (SHAP=0,312) merupakan faktor mediator utama dalam mekanisme dampak MBG.
- Fitur unduh PDF dan Excel memungkinkan distribusi hasil analisis kepada pemangku kepentingan secara efisien.

## **BAB V**

### **PENUTUP**

#### **5.1 Kesimpulan**

Berdasarkan hasil implementasi Dashboard Analisis Dampak Akademik Program Makan Bergizi Gratis (MBG) dan analisis yang telah dilakukan terhadap data 270 siswa dari 6 Sekolah Dasar di Kecamatan Binjai Utara, Kota Binjai, Sumatera Utara, dapat ditarik kesimpulan sebagai berikut:

##### **5.1.1 Dampak Signifikan MBG terhadap Prestasi Akademik**

Program Makan Bergizi Gratis (MBG) terbukti memberikan dampak yang positif dan signifikan secara statistik terhadap prestasi akademik siswa SD di Kecamatan Binjai Utara. Analisis Difference-in-Differences (DiD) menghasilkan Average Treatment Effect (ATE) keseluruhan sebesar +0,68 poin ( $p = 0,000$ ; CI 95%: [0,41; 1,00]), yang berarti siswa penerima MBG mengalami peningkatan nilai rata-rata 0,68 poin lebih tinggi dibandingkan kelompok kontrol setelah dikontrol melalui Propensity Score Matching.

Dampak MBG terhadap mata pelajaran Matematika (ATE = +0,70;  $p = 0,000$ ; CI: [0,41; 1,00]) sedikit lebih besar dibandingkan terhadap Bahasa Indonesia (ATE = +0,66;  $p = 0,000$ ; CI: [0,37; 0,96]). Perbedaan ini mengindikasikan bahwa perbaikan status gizi melalui program MBG lebih besar dampaknya terhadap kemampuan analitis dan berpikir logis (Matematika) dibandingkan kemampuan berbasis literasi (Bahasa Indonesia), sesuai dengan mekanisme neurologis yang menghubungkan gizi dengan fungsi kognitif eksekutif.

### **5.1.2 Kebersihan Keseimbangan Kolompok melalui PSM**

Proses Propensity Score Matching berhasil menyeimbangkan kovariat antara kelompok treatment dan kontrol dengan baik. Dari 270 siswa dalam dataset, 108 pasang berhasil dicocokkan. Standardized Mean Difference (SMD) rata-rata turun dari 0,307 menjadi 0,094 setelah matching, dengan 5 dari 6 variabel kovariat mencapai status Seimbang ( $SMD < 0,1$ ). Variabel kehadiran memiliki SMD setelah matching 0,168 (status Cukup) namun telah mengalami reduksi 78% dari nilai awal. Keberhasilan PSM memvalidasi bahwa estimasi DiD bebas dari bias seleksi yang substansial, sehingga kesimpulan kausal mengenai efek MBG dapat dipertanggungjawabkan secara metodologis.

### **5.1.3 Performa Model Gradient Boosting dalam Klasifikasi Dampak**

Model Gradient Boosting Classifier berhasil diimplementasikan untuk mengklasifikasikan perubahan nilai siswa ke dalam tiga kategori (Meningkat, Stabil, Menurun) dengan akurasi 66,7% dan F1-Score macro 0,655 pada data testing. Model menunjukkan performa terbaik pada kelas Meningkatkan (29 dari 35 prediksi benar), sementara kelas Stabil dan Menurun memiliki performa lebih rendah yang kemungkinan disebabkan oleh ketidakseimbangan distribusi kelas. Ketika diterapkan pada seluruh dataset, model mengklasifikasikan 176 siswa (65,2%) sebagai Meningkatkan, 70 siswa (25,9%) sebagai Stabil, dan 24 siswa (8,9%) sebagai Menurun, dengan rata-rata kepercayaan prediksi 88,2%.

### **5.1.4 Faktor-Faktor yang Mempengaruhi Dampak MBG**

Analisis SHAP mengidentifikasi Kehadiran Sekolah sebagai faktor paling berpengaruh terhadap dampak MBG (Mean  $|SHAP| \approx 0,28$ ), diikuti oleh Status MBG itu sendiri ( $\approx 0,21$ ), nilai Pre-Bahasa Indonesia ( $\approx 0,18$ ), dan nilai Pre-

Matematika ( $\approx 0,14$ ). Temuan ini mengindikasikan bahwa program MBG bekerja melalui dua mekanisme utama: (1) mekanisme tidak langsung melalui peningkatan kehadiran sekolah—siswa yang mendapat makan bergizi lebih termotivasi dan mampu hadir di sekolah; dan (2) mekanisme langsung melalui peningkatan kapasitas kognitif yang tercermin dari nilai SHAP Status MBG di peringkat kedua.

### **5.1.5 Keberhasilan Implementasi Dashboard**

Dashboard Analisis Dampak Akademik MBG berhasil diimplementasikan sebagai aplikasi web full-stack yang mengintegrasikan analisis statistik lanjutan (PSM, DiD), machine learning (Gradient Boosting), dan interpretabilitas model (SHAP) dalam satu antarmuka yang interaktif dan ramah pengguna. Dashboard mencakup sistem autentikasi JWT dengan tiga peran (Admin, Peneliti, Viewer), visualisasi data yang responsif menggunakan Recharts, pipeline analisis otomatis, fitur prediksi individual, segmentasi siswa, dan ekspor hasil analisis dalam format PDF dan Excel. Arsitektur berbasis React + FastAPI dengan desain API RESTful memungkinkan skalabilitas sistem untuk dataset yang lebih besar dan cakupan sekolah yang lebih luas.

## **5.2 Saran**

Berdasarkan temuan penelitian dan keterbatasan yang telah diidentifikasi, berikut adalah saran-saran yang ditujukan kepada berbagai pemangku kepentingan:

### **5.2.1 Saran untuk Implementasi Program MBG**

1. Meningkatkan konsistensi pelaksanaan program MBG, terutama dalam hal frekuensi dan kualitas pemberian makanan. Mengingat kehadiran

sekolah merupakan mediator utama dampak MBG, pihak sekolah perlu memastikan bahwa program MBG berjalan secara rutin untuk mempertahankan efek motivasi kehadiran.

2. Memberikan perhatian khusus kepada 24 siswa (8,9%) yang diprediksi mengalami penurunan nilai dan 24 siswa yang dikategorikan sebagai risiko tinggi. Intervensi tambahan seperti bimbingan belajar, konseling, atau dukungan gizi tambahan perlu dipertimbangkan untuk kelompok ini.
3. Memperluas program MBG kepada sekolah-sekolah yang belum terjangkau di Kecamatan Binjai Utara, mengingat dampak positif yang telah terbukti signifikan secara statistik.

#### **5.2.2 Saran untuk Pengembangan Penelitian Selanjutnya**

1. Menggunakan data riil dari lapangan (bukan data simulasi) untuk memvalidasi dan memperkuat temuan penelitian ini. Data riil akan memberikan estimasi yang lebih akurat mengenai dampak nyata program MBG terhadap prestasi akademik.
2. Memperluas variabel outcome untuk mencakup indikator-indikator lain seperti tingkat konsentrasi belajar, tingkat kehadiran aktual, dan indikator kesehatan (berat badan, tinggi badan, hemoglobin) untuk mendapatkan gambaran dampak MBG yang lebih komprehensif.
3. Menambahkan analisis longitudinal dengan lebih dari dua titik pengukuran waktu untuk memahami dinamika dampak MBG dalam jangka panjang dan mengidentifikasi titik-titik kritis dalam mekanisme dampak.
4. Melakukan analisis heterogenitas dampak untuk mengidentifikasi

subkelompok siswa yang paling diuntungkan oleh program MBG berdasarkan karakteristik seperti status sosial ekonomi, kelas, atau kondisi kesehatan awal.

5. Mengoptimalkan model Gradient Boosting menggunakan teknik hyperparameter tuning (GridSearchCV atau RandomizedSearchCV) dan class balancing (SMOTE atau class\_weight) untuk meningkatkan performa prediksi terutama pada kelas minoritas (Stabil dan Menurun).

### **5.2.3 Saran untuk Pengembangan Dashboard**

1. Mengintegrasikan koneksi langsung ke database sekolah untuk memungkinkan pembaruan data secara real-time tanpa perlu input manual, meningkatkan efisiensi dan akurasi analisis.
2. Menambahkan fitur notifikasi otomatis kepada guru atau orang tua siswa yang diidentifikasi berisiko tinggi oleh model prediksi.
3. Mengimplementasikan dashboard mobile-responsive atau aplikasi mobile khusus agar dapat diakses oleh kepala sekolah dan guru dengan lebih mudah dalam kegiatan lapangan.
4. Mempertimbangkan deployment ke infrastruktur cloud yang lebih permanen (AWS, GCP, atau Azure) untuk memastikan ketersediaan sistem yang tinggi dan skalabilitas dalam menghadapi jumlah sekolah dan siswa yang lebih besar.

## DAFTAR PUSTAKA

- Abukader, A., Alzubi, J., & Adegboye, O. (2025). Student performance prediction using Light Gradient Boosting Machine optimized with metaheuristic algorithms and SHAP interpretability. *Journal of Educational Data Mining*.
- Amini-Rarani, M., et al. (2024). Quasi-experimental designs for policy evaluation in education. *Journal of Educational Policy*, 39(2), 145–167.
- Appiah, E. (2024). Impact of school feeding programs on student attendance and academic performance in rural Ghana. *African Journal of Educational Studies*, 22(1), 45–62.
- Arifin, M., et al. (2025). Dampak program makan bergizi terhadap konsentrasi dan prestasi akademik siswa di Banda Aceh. *Jurnal Pendidikan Indonesia*, 14(1), 78–92.
- Azis, A. (2024). Perbandingan algoritma machine learning untuk prediksi performa akademik mahasiswa. *Jurnal Teknologi Informasi*, 12(3), 210–225.
- Bracic, A., et al. (2022). Propensity score matching in educational quasi-experiments. *Educational Evaluation and Policy Analysis*, 44(3), 401–420.
- Carlisle, D. M., et al. (2023). Quasi-experimental designs in public health and education policy evaluation. *American Journal of Public Health*, 113(4), 512–520.
- Durrant, G. B. (2025). Propensity score matching: A practical guide for social scientists. *SAGE Research Methods*.
- Hoddinott, J., et al. (2021). Early childhood nutrition interventions and cognitive development: Long-term evidence. *The Lancet Global Health*, 9(5), e621–e629.
- Melkamu, A., et al. (2022). Propensity score matching for policy impact evaluation in developing countries. *World Development*, 151, 105742.

- Molnar, C. (2022). *Interpretable machine learning: A guide for making black box models explainable* (2nd ed.). <https://christophm.github.io/interpretable-ml-book/>
- Munandar, R., et al. (2025). Peran multidimensional program makan bergizi gratis dalam mendukung kualitas pembelajaran. *Journal of Educational Management Research*, 8(1), 112–129.
- Özkurt, S. (2024). Comparative analysis of machine learning algorithms in educational data prediction. *International Journal of Artificial Intelligence in Education*, 34(2), 301–325.
- Rizki, F., & Darmayanti, R. (2025). Kebijakan berbasis data untuk peningkatan hasil belajar di Indonesia. *Jurnal Kebijakan Pendidikan*, 17(1), 33–49.
- Rizkallah, J. (2025). Gradient Boosting Trees for academic performance prediction. *Journal of Machine Learning Research*, 26, 1–25.
- Sinaga, R., & Syarief, A. (2025). Variasi dampak program makan bergizi akibat kendala logistik di Indonesia. *Jurnal Gizi dan Pangan*, 20(1), 55–70.
- Solihat, N., et al. (2025). Dampak program makan bergizi terhadap motivasi belajar dan partisipasi siswa. *Jurnal Psikologi Pendidikan*, 13(2), 144–159.
- Wang, L., et al. (2021). School feeding programs in developing countries: A systematic review and meta-analysis. *The American Journal of Clinical Nutrition*, 113(4), 987–1001.

- Wall, T., et al. (2022). Comparative study of school feeding programs across African nations. *International Journal of Educational Development, 91*, 102587.
- Austin, P. C. (2011). An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behavioral Research, 46*(3), 399–424.  
<https://doi.org/10.1080/00273171.2011.568786>
- Bautista-Jiménez, E., & Flores-Alfaro, E. (2017). School feeding programs and academic performance: A systematic review. *Salud Pública de México, 59*(3), 321–332.
- Che Rossb, M. S., Mohd Azizan, N., & Zainuddin, A. (2019). The effect of school breakfast program on academic performance among primary school students. *Malaysian Journal of Medicine and Health Sciences, 15*(3), 56–63.
- Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 785–794*.  
<https://doi.org/10.1145/2939672.2939785>
- Dahl, W. J., & Stewart, M. L. (2015). Position of the Academy of Nutrition and Dietetics: Health implications of dietary fiber. *Journal of the Academy of Nutrition and Dietetics, 115*(11), 1861–1870.
- Dickson, R., Awasthi, S., Williamson, P., Demellweek, C., & Garner, P. (2000). Effects of treatment for intestinal helminth infection on growth

- and cognitive performance in children: Systematic review of randomized trials. *BMJ*, 320(7251), 1697–1701.
- Fanelli Kuczmariski, M., & Mason, M. A. (2016). Nutrition and aging. In *Practical approaches to geriatric care* (pp. 89–108). Springer.
- Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *The Annals of Statistics*, 29(5), 1189–1232.
- Griggs, J. L., Causbie, M., & Turner, S. (2017). Food insecurity and academic performance: Systematic review and meta-analysis. *Comprehensive Reviews in Food Science and Food Safety*, 16(3), 455–466.
- Heckman, J. J., Ichimura, H., & Todd, P. E. (1997). Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme. *Review of Economic Studies*, 64(4), 605–654.
- Ho, D. E., Imai, K., King, G., & Stuart, E. A. (2011). MatchIt: Nonparametric preprocessing for parametric causal inference. *Journal of Statistical Software*, 42(8), 1–28.
- Holford, P. (2019). *Optimum Nutrition for the Mind* (3rd ed.). Piatkus Books.
- Imai, K., & Ratkovic, M. (2014). Covariate balancing propensity score. *Journal of the Royal Statistical Society: Series B*, 76(1), 243–263.
- Jomaa, L. H., McDonnell, E., & Probart, C. (2011). School feeding programs in developing countries: Impacts on children's health and educational outcomes. *Nutrition Reviews*, 69(2), 83–98. <https://doi.org/10.1111/j.1753-4887.2010.00369.x>

- Jukes, M. C. H., Drake, L. J., & Bundy, D. A. P. (2008). School health, nutrition and education for all: Levelling the playing field. *CAB International*.
- Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi Republik Indonesia. (2022). *Pedoman pelaksanaan program makan bergizi di sekolah*. Kemendikbudristek.
- Kementerian Kesehatan Republik Indonesia. (2019). *Peraturan Menteri Kesehatan Nomor 28 Tahun 2019 tentang Angka Kecukupan Gizi yang Dianjurkan untuk Masyarakat Indonesia*. Kemenkes RI.
- Kleinman, R. E., Hall, S., Green, H., Korzec-Ramirez, D., Patton, K., Pagano, M. E., & Murphy, J. M. (2002). Diet, breakfast, and academic performance in children. *Annals of Nutrition and Metabolism*, 46(1), 24–30.
- Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, 30, 4765–4774.
- Magnuson, K., & Waldfogel, J. (2016). Trends in income-related gaps in readiness for school: 1998 to 2010. *AERA Open*, 2(3), 1–17.
- Nasution, A. S., & Daulay, N. (2023). Implementasi program gizi sekolah terhadap peningkatan prestasi belajar siswa SD di Sumatera Utara. *Jurnal Pendidikan dan Kesehatan Masyarakat*, 10(2), 78–89.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12, 2825–2830.

- Peraturan Presiden Republik Indonesia Nomor 83 Tahun 2017 tentang Kebijakan Strategis Pangan dan Gizi. (2017). Sekretariat Negara.
- Pollit, E. (1995). Does breakfast make a difference in school? *Journal of the American Dietetic Association*, 95(10), 1134–1139.
- Prasetyo, B., & Kurniawan, R. (2022). Analisis dampak program pemberian makanan bergizi terhadap nilai ujian siswa sekolah dasar menggunakan metode Difference- in-Differences. *Jurnal Teknologi Informasi dan Pendidikan*, 15(1), 45–57.
- Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1), 41–55. <https://doi.org/10.1093/biomet/70.1.41>
- Rubin, D. B. (2001). Using propensity scores to help design observational studies: Application to the tobacco litigation. *Health Services & Outcomes Research Methodology*, 2, 169–188.
- Sari, D. K., & Pratiwi, A. (2021). Hubungan status gizi dengan prestasi belajar anak sekolah dasar di Kota Binjai. *Jurnal Gizi dan Kesehatan*, 13(2), 112–120.
- Shapley, L. S. (1953). A value for n-person games. In H. Kuhn & A. Tucker (Eds.), *Contributions to the Theory of Games* (Vol. 2, pp. 307–317). Princeton University Press.
- Stuart, E. A. (2010). Matching methods for causal inference: A review and a look forward. *Statistical Science*, 25(1), 1–21.
- Taras, H. (2005). Nutrition and student performance at school. *Journal of School Health*, 75(6), 199–213.

UNICEF. (2020). *The state of the world's children 2019: Children, food and nutrition*. United Nations Children's Fund.

World Food Programme. (2021). *State of school feeding worldwide 2020*.  
World Food Programme.

Yusuf, A. M. (2014). *Metode penelitian kuantitatif, kualitatif, dan penelitian gabungan*. Prenadamedia Group.

Zhao, Q., & Percival, D. (2017). Entropy balancing is doubly robust. *Journal of Causal Inference*, 5(1), 1–15.

