

Menulis Jurnal di Bantu AI: Data-First Anti-Halusinasi

Onno W. Purbo
onno@indo.net.id
onno@itts.ac.id
Twitter onnowpurbo

Institut Teknologi Tangerang Selatan (ITTS)
OnnoCenter
2025

Lisensi & Catatan Karya

Karya ini dilisensikan di bawah lisensi **Creative Commons Attribution-ShareAlike 4.0 International (CC BY-SA 4.0)**.

Artinya, **siapa pun bebas untuk menggunakan, menyalin, membagikan, dan mengadaptasi** materi ini, **dengan syarat**:

- **Memberikan atribusi yang sesuai** (menyebut sumber asli),
- **Menyertakan lisensi yang sama** jika dimodifikasi atau dikembangkan lebih lanjut.

Lihat detail lisensi di sini <https://creativecommons.org/licenses/by-nd/4.0/legalcode.id>

Disclaimer

Buku ini disusun untuk tujuan **edukasi dan pengembangan kapasitas akademik**, bukan sebagai panduan resmi institusi atau lembaga penerbit jurnal mana pun. Seluruh contoh, *prompt*, dan ilustrasi bersifat demonstratif untuk membantu mahasiswa dan peneliti memahami prinsip **Data-First** dan praktik **penulisan ilmiah berbasis bukti** di era AI. Penulis tidak bertanggung jawab atas penyalahgunaan alat atau hasil AI yang melanggar etika akademik. Segala bentuk analisis, data, dan rekomendasi dalam buku ini harus tetap diverifikasi secara mandiri sebelum digunakan dalam konteks penelitian atau publikasi resmi.

Daftar Isi

Disclaimer.....	2
Daftar Isi.....	3
Kata Pengantar.....	5
Abstract.....	6
Catatan Praktis.....	7
Bagian 0 — Manifesto Data-First.....	8
Mengapa Data adalah Segalanya.....	8
Manifesto Data-First.....	8
Urutan Kerja yang Harus Dihindari.....	11
Urutan Kerja yang Lebih Ampuh (Data-First).....	14
Bagian I — Menyiapkan Bahan: Eksperimen, Observasi, dan Data.....	17
Merancang Eksperimen & Observasi yang “Ramah AI”.....	17
Kebersihan Data & Dokumentasi.....	21
Etika & Kepatuhan: Pondasi Kepercayaan Ilmiah.....	24
Bagian II — Dari Data ke Metodologi & Hasil.....	27
Analisis Data Inti (EDA Cepat & Tajam).....	27
Menulis Metode yang Reproducible (Dibantu AI).....	31
Menulis Hasil: Angka Dulu, Tafsir Sesudah.....	35
Bagian III — Menemukan & Memastikan Novelty.....	39
Definisi Novelty yang Terukur.....	39
Cek Novelty dengan AI (Anti-Halusinasi).....	43
Kotak Prompt — Paket Siap Pakai.....	45
Do & Don’t saat cek novelty dengan AI.....	46
Ringkasan kilat.....	46
Jika Novelty Lemah: Iterasi Eksperimen Terarah.....	47
Bagian IV — Mengkuantifikasi Klaim (Dari “A lebih baik” → “A 56% lebih baik”).....	50
Metrik & Rumus Perbandingan.....	50
Relative Improvement.....	51
Membuat Pernyataan Terukur yang Kuat.....	54
Visual Kuantitatif yang Meyakinkan.....	57
Bagian V — Menyusun Kesimpulan → Abstrak → Pendahuluan (Urutan Terbalik yang Efisien).....	61
Kesimpulan Terarah Aksi (Actionable).....	61
Tantangan singkat.....	63
Abstrak yang Tajam & Critical (Highlight Novelty).....	64
Pendahuluan Tanpa Halusinasi.....	68
Bagian VI — Tinjauan Pustaka & Perbandingan dengan Studi Lain.....	72
Mencari Referensi yang Tidak Halusinasi.....	72
Membandingkan Hasil Kita dengan Literatur.....	75
Menulis Related Work yang “Nendang”.....	78
Bagian VII — Menentukan Judul Terakhir.....	81

Judul yang Meng-highlight Novelty.....	81
Template Praktis (tinggal isi).....	81
Bagian VIII — Manajemen Referensi & Kepatuhan.....	84
Zotero/Mendeley/EndNote: Workflow Anti-Ribet.....	84
Transparansi & Replikasi.....	88
Bagian IX — Siap Kirim & Menjawab Reviewer.....	92
Memilih Jurnal & Menyesuaikan Naskah.....	92
Checklist Pra-Submission (Satu Halaman).....	95
Strategi Menjawab Reviewer dengan Data.....	99
Lampiran A: Template Tabel Komparasi Kuantitatif (siap isi).....	103
Lampiran B: Kamus Metrik & Rumus (efek, CI, uji cepat).....	106
Lampiran C: Kumpulan Prompt Siap Pakai.....	110
Glossary.....	115
Referensi.....	120
Tentang Penulis.....	121

Kata Pengantar

Alhamdulillah, Puji syukur penulis panjatkan kepada Allah SWT atas tersusunnya buku ***Menulis Jurnal di Bantu AI: Data-First Anti-Halusinasi***. Buku ini lahir dari kebutuhan nyata di dunia akademik: bagaimana menggunakan **kecerdasan buatan secara cerdas dan etis** untuk menulis karya ilmiah yang **berbasis data, bebas halusinasi, dan mudah diaudit**.

Dalam beberapa tahun terakhir, banyak mahasiswa dan peneliti yang terjebak pada *shortcut*—mengandalkan AI tanpa data, tanpa verifikasi, dan tanpa jejak audit. Akibatnya, tulisan tampak ilmiah, tetapi kosong dari bukti. Buku ini ingin membalikkan kebiasaan itu dengan pendekatan ***Data-First***, yaitu menulis hanya setelah angka, tabel, dan grafik siap. **Narasi mengikuti data, bukan sebaliknya.**

Melalui buku ini, mahasiswa diajak untuk:

- **Berpikir seperti ilmuwan sejati:** setiap klaim harus punya angka pendukung.
- **Menggunakan AI sebagai asisten, bukan pengganti riset.**
- **Membangun kebiasaan audit trail,** agar semua hasil bisa diverifikasi.
- **Mengembangkan *prompt engineering* yang bertanggung jawab,** agar AI membantu, bukan menyesatkan.

Pendekatan *Data-First* juga memperkenalkan konsep *anti-halusinasi*—yakni setiap kalimat ilmiah harus bisa ditelusuri ke data mentah, tabel, atau referensi yang sah. Dengan begitu, mahasiswa belajar **menulis dengan integritas**, bukan sekadar menyusun kata-kata indah.

Penulis berharap buku ini dapat menjadi **panduan praktis bagi mahasiswa, dosen, dan peneliti** yang ingin menulis jurnal dengan cara yang lebih modern, efisien, dan etis. Jadikan AI teman belajar, bukan jalan pintas. Karena di era digital ini, **yang membedakan peneliti unggul bukan siapa yang paling cepat menulis, tetapi siapa yang paling bisa membuktikan datanya.**

Selamat bereksperimen, belajar, dan menulis dengan semangat *open science*.

Jakarta, October 2025

Onno W. Purbo

Abstract

Buku ini memperkenalkan pendekatan **Data-First Anti-Halusinasi** sebagai paradigma baru dalam penulisan jurnal ilmiah di era kecerdasan buatan. Di tengah maraknya penggunaan *AI writing tools* yang sering menghasilkan teks tanpa dasar data, buku ini menegaskan bahwa **data adalah fondasi utama keilmuan**, bukan sekadar pelengkap narasi. Pendekatan *Data-First* menuntut peneliti untuk **memulai dari angka, tabel, dan grafik**, kemudian membangun narasi yang dapat diverifikasi dan diaudit. Dengan strategi ini, mahasiswa dan dosen diarahkan untuk menulis secara **transparan, terukur, dan reproducible**, sambil tetap memanfaatkan AI secara cerdas sebagai asisten analitik dan literasi ilmiah. Buku ini juga membahas konsep **anti-halusinasi ilmiah**, yaitu disiplin memastikan setiap klaim memiliki rujukan data atau sumber sah yang dapat dilacak. Melalui panduan, contoh *prompt*, dan template siap pakai, pembaca diajak memahami cara mengintegrasikan AI dalam proses ilmiah tanpa kehilangan integritas akademik. Akhirnya, karya ini menjadi **manifesto menulis dengan bukti**, bukan sekadar opini — menempatkan peneliti muda Indonesia pada jalur yang lebih *evidence-driven*, etis, dan kompetitif di dunia riset global.

Catatan Praktis

- **Mulai selalu dari data:** tanpa tabel/grafik/angka, jangan menulis narasi.
- Setiap klaim komparatif **wajib** angka (sertakan metrik & kondisi uji).
- Targetkan minimal **tiga novelty** yang dapat diukur/divalidasi.
- Pendahuluan & *literature review* **ditulis setelah** angka & novelty solid.
- Simpan semua *prompt–output–sumber* sebagai *appendix audit* untuk integritas ilmiah.

EDA, deteksi tren/anomali, kuantifikasi klaim, cek novelty, abstrak, pendahuluan, *related work*, respond reviewer

A.. **Checklist Anti-Halusinasi Referensi** (DOI, penerbit, bukti halaman/kutipan)

B. **Template Abstrak 150–250 kata (2 gaya)**

C. **Template Response to Reviewers** (tabel: komentar → respons → bukti)

D. **Flowchart Proses Data-First** (Data → Analisis → Novelty → Kesimpulan → Abstrak → Pendahuluan → Judul)

Bagian 0 — Manifesto Data-First

Mengapa Data adalah Segalanya

Manifesto Data-First

Mengapa Data adalah Segalanya

Bayangkan menulis tanpa data. Itu seperti bikin peta tanpa kompas. Serba ngawur. **Data adalah bahan baku utama.** Tanpa data, argumen hanyalah opini.

- **Data memberi bukti.** Bukan sekadar “katanya”.
- **Data menahan bias.** Kita diuji oleh angka, bukan perasaan.
- **Data membuat ulang bisa dilakukan.** Orang lain dapat mengulang dan memverifikasi.

Ingat: jurnal ilmiah itu *evidence-based*. **Tanpa bukti, naskahmu rapuh.**

Mini tantangan mahasiswa:

Ambil topik yang kamu suka (misal: *study habit*). Tulis 1 klaim **tanpa** data. Lalu tulis ulang klaim **dengan** angka sederhana (n , rata-rata, selisih). Rasakan bedanya!

Data sebagai syarat sah menulis naskah

Menulis naskah = mempresentasikan bukti. Jadi, **data adalah “tiket masuk”**.

- **Minimal harus ada:**
 - **Data mentah** (*raw data*): tabel nilai, log eksperimen, hasil survei.
 - **Instrumen & prosedur:** bagaimana data dikumpulkan.
 - **Ringkas analisis:** rata-rata, deviasi, uji sederhana.
- **Tanpa data, tolak diri sendiri.** Serious.
- **Contoh kalimat sah:**
 - “Metode A mencapai akurasi **82.4%** pada Dataset X ($n=500$).”
 - “Metode B **63% lebih cepat** (waktu 18s vs 49s; $\Delta=31s$).”

Checklist cepat (boleh tempel di laptop):

- Punya **tabel** angka pokok
- Punya **gambar** atau grafik yang jelas
- Punya catatan singkat **metode** pengambilan data
- Punya **file data** rapi & bisa dibagikan (CSV/Parquet)

AI = akselerator, bukan pengganti riset

AI hebat untuk **mempercepat**. Bukan untuk **mengada-ada**.

- **Yang boleh AI bantu:** merapikan kalimat, merangkum, menyarankan analisis, bikin *draft* awal.
- **Yang tidak boleh AI ganti:** eksperimen/pengamatan, pengukuran, kejujuran ilmiah.
- **Bahaya utama:** *halusinasi*. AI bisa “PD” bikin referensi palsu atau angka asumsi.

Cara aman pakai AI:

- **Masukkan data kamu** (tabel/gambar) → minta AI **membaca** dan **menjelaskan**.
- **Minta angka spesifik** (misal *relative improvement*).
- **Verifikasi manual** semua output AI.

Prompt contoh (aman & berguna):

“Dari **Tabel 1** ini, hitungkan selisih A vs B, *relative improvement*, dan tulis satu kalimat hasil.”

“Periksa kalimat ini: ‘A lebih baik dari B’. Ubah jadi versi **kuantitatif** dengan angka dari **Gambar 2**.”

Rule of thumb: AI membantu menulis. Data memegang kendali.

Traceability & audit trail — tiap klaim kembali ke tabel/gambar

Setiap klaim harus bisa ditelusuri kembali ke **sumber angka**. Ini namanya **traceability**. Rekaman proses dari data mentah → olah → angka akhir = **audit trail**.

- **Klaim → Tabel/Gambar → Data mentah.**
- Beri **label jelas**: “Lihat **Tabel 2** baris 4”, atau “lihat **Gambar 3** panel (b)”.
- Simpan *notebook* atau *script* analisis. Catat versi alat (*versioning*).
- Buat **folder rapi**: */data_raw/*, */data_clean/*, */figures/*, */analysis/*.

Contoh kalimat ber-*traceability*:

- “Metode A **29% lebih akurat** dari B pada Dataset X (**Tabel 2**, kolom *Accuracy*).”
- “Waktu inferensi turun **56%** (49s → 21.5s; **Gambar 4**, panel (a)).”

Mini praktik 10 menit:

1. Pilih **satu klaim** dari tugas/proyekmu.
2. Tambahkan **referensi balik** ke tabel/gambar (nomor & lokasi).
3. Simpan **file sumber angka** di folder **/data_clean/**, dan catat di README 3 baris:
 - Sumber data
 - Tanggal olah
 - Alat/versi (*software, library*)

Goal-nya: pembaca bisa **klik–lihat–cek**. **Transparan. Reproducible.**
Di *peer review*, ini yang bikin naskahmu **percaya diri**.

Ringkasnya (buat ditempel di meja belajar)

- **Data dulu.** Narasi belakangan.
- **Naskah sah** = ada **data, metode, angka**.
- **AI membantu ngebut, kamu yang bertanggung jawab.**
- **Setiap klaim harus bisa ditelusuri ke tabel/gambar.**
- **Tanpa data → no paper.** Dengan data → **paper kuat dan siap diuji.**

Semangat, kawan. Jadikan **angka** teman terbaikmu. Jadikan **AI** sopir cadanganmu. Kamu tetap pegang setirnya.

Urutan Kerja yang Harus Dihindari

Urutan Kerja yang Harus Dihindari

Kamu pasti sering dengar saran klasik saat dosen bilang:

“Mulailah menulis dari **judul**, lanjut ke **abstrak**, lalu **pendahuluan**, **metode**, **hasil**, **kesimpulan**, dan terakhir **referensi**.”

Kelihatannya rapi dan logis, kan? Tapi sayangnya, **itu jebakan klasik**.

Pola ini bikin banyak mahasiswa **stuck di halaman satu** selama berminggu-minggu.

Kenapa Pola Klasik Ini Menyusahkan?

Mari kita bongkar satu per satu dengan cara santai.

1. **Mulai dari judul = macet di awal.**

Karena kamu belum punya data, belum tahu hasil, tapi sudah dipaksa menentukan “tema besar”. Akibatnya, judul sering berubah-ubah, dan kamu kehilangan arah.

→ *Ibarat mau masak tapi belum tahu bahan, tapi udah sibuk mikirin nama makanannya.*

2. **Abstrak terlalu dini = halu akademik.**

Abstrak itu seharusnya ringkasan dari *semua hasil penelitian*, bukan ramalan masa depan. Kalau kamu menulis abstrak sebelum punya hasil, yang kamu tulis adalah **fiksi ilmiah**.

3. **Pendahuluan dulu = capek sebelum berlari.**

Ini bagian paling berat, karena harus mencari referensi, gap riset, dan teori. Kalau kamu belum punya data, menulis pendahuluan jadi *ngawang-ngawang*. Otakmu dipaksa menebak masalah yang bahkan belum kamu buktikan.

4. **Metode tanpa data = template kosong.**

Menulis metode di awal tanpa eksperimen nyata sering berakhir seperti *copy-paste* dari paper lain. Hasilnya? Tidak orisinal, dan susah disesuaikan saat data sudah datang.

5. **Hasil terakhir = stres level dewa.**

Ketika semua bagian lain sudah selesai, barulah kamu menulis hasil. Tapi ternyata hasilnya tidak sesuai hipotesis. Akhirnya, seluruh tulisan sebelumnya **tidak nyambung**. Kamu harus revisi besar-besaran.

6. **Kesimpulan & referensi jadi formalitas.**

Karena sudah lelah di akhir, kesimpulan hanya diisi kalimat umum seperti “Penelitian ini diharapkan dapat bermanfaat bagi masyarakat.”

Tidak ada angka, tidak ada bukti, tidak ada *impact*.

Akibat Nyata dari Pola Ini

Pola klasik ini bikin proses penulisan jadi:

- **Lambat**, karena kamu menulis dari bagian paling abstrak ke paling konkret.
- **Melelahkan**, karena bolak-balik revisi akibat tidak punya data awal.
- **Berisiko halusinasi**, karena kamu terpaksa menulis hal yang belum terbukti.
- **Tidak fokus**, karena ide dan data berjalan di dua dunia yang berbeda.

Bahkan dengan bantuan *AI*, urutan ini tetap berbahaya. *AI* hanya bisa mempercepat kalau kamu punya bahan — **bukan mengarang data!**

Analogi Sederhana

Menulis jurnal pakai urutan lama itu seperti:

“Bangun rumah mulai dari atap.”

Kamu belum punya pondasi (data), belum punya tiang (metode), tapi sudah sibuk menghias genteng (judul dan abstrak). Akhirnya rumahmu roboh begitu diuji reviewer.

Kuncinya: Balik Pola!

Mulailah dari **data**, bukan dari **judul**. Dari **bukti nyata**, bukan dari **narasi spekulatif**.

- Eksperimen dulu.
- Lihat hasilnya.
- Baru tentukan cerita ilmiahnya.

Menulis jurnal itu seperti detektif:

kumpulkan bukti dulu, baru buat cerita yang masuk akal. Jangan menulis cerita dulu baru cari bukti yang cocok. Itu namanya **fit data** — bukan riset!

Catatan Buat Kamu

- **Jangan takut “melanggar tradisi”**. Pola klasik itu untuk abad 20. Di era *AI* dan *data-first*, yang penting adalah **isi dan bukti**, bukan urutan formal.
- **AI boleh bantu nulis, tapi kamu harus punya dasar data**. Tanpa itu, *AI* akan “mengarang dengan sopan”.
- **Reviewer menghormati angka, bukan opini**.

Tantangan Kecil untuk Mahasiswa:

Coba buka tugas akhir atau proyek risetmu sekarang. Lihat apakah kamu sudah:

- Menulis **judul** sebelum punya hasil?
- Menyusun **abstrak** tanpa data?
- Membuat **pendahuluan** sebelum eksperimen?

Kalau iya, coba hentikan dulu. Mulailah kumpulkan **tabel, grafik, hasil observasi**, sekecil apa pun. Lalu baru gunakan *AI* untuk membantu menulis berdasarkan data itu.

Ingat mantra kita:

“Data dulu, tulisan belakangan.”

Dengan pendekatan ini, kamu tidak hanya menulis lebih cepat, tapi juga lebih **jujur, akurat, dan keren secara akademik**.

Urutan Kerja yang Lebih Ampuh (Data-First)

(1) Eksperimen / Pengamatan / Kumpulan Data

Mulai dari **dunia nyata**, bukan dari imajinasi judul. Kumpulkan **bukti**: angka, tabel, log, foto, grafik, rekaman alat.

- **Tips cepat:**
 - Buat **sheet metadata** (kapan, di mana, alat, versi).
 - Simpan data rapi: **/data_raw/**, **/data_clean/**, **/figures/**.
 - Tulis langkah singkat di *lab notebook* digital (boleh Google Docs/Notion).
- **Mini tantangan:**
Hari ini, kumpulkan **10 baris data** apa pun (belajar 1 jam/hari, detak jantung lari, waktu loading aplikasi). Besok, tambah 10 lagi. Konsisten itu penting.
- **Prompt aman untuk AI:**

*“Bacalah **Tabel 1** (10 baris). Apa pola paling terlihat? Sebutkan 3 dugaan awal dalam 1 kalimat per dugaan.”*

(2) Analisis Data, Metode, Hasil

Lakukan *EDA* sederhana dulu. Lihat **tren**, **outlier**, **perbandingan**.

- **Langkah ringkas:**
 - Cek **rata-rata/median**, **sebaran**, **grafik garis/batang**.
 - Catat **alat & parameter** (*software*, *library*, versi).
 - Simpan **gambar** yang enak dilihat (label jelas, satu pesan per grafik).
- **Format paragraf hasil yang mudah:**
 - **Angka dulu** → **bandingkan** → **makna awal**.
Contoh: “Metode A akurasi **82.4%**, B **63.9%** ($\Delta=+18.5$ pts). Ini menunjukkan A lebih stabil pada data bising.”
- **Prompt praktis:**

“Dari grafik ini, tulis 3 kalimat hasil: (1) angka utama, (2) selisih vs baseline, (3) kemungkinan penyebab.”

(3) Pastikan ≥ 3 Novelty yang Jelas

Target minimal **tiga hal baru** yang bisa diuji ulang. Jangan samar.

- **Jenis *novelty* yang sah:**
 - **Metode baru/variasi** (mis. *preprocessing* unik).
 - **Domain/data baru** (konteks Indonesia, perangkat murah).
 - **Performa/efisiensi** (lebih cepat, lebih hemat, *robust*).
 - **Insight/interpretasi** yang bikin paham “*kenapa*”.
- **Checklist *novelty* (pilih 3+):**
 - Δ akurasi/MAE/F1 **jelas** dan **terukur**.
 - Penghematan waktu/biaya/energi.
 - Generalisasi di sub-kelompok (mis. data minoritas).

- **Prompt cek *novelty*:**

“Dari data yang di attached, jelaskan novelty apa yang bisa kita peroleh”

*“Bandingkan hasil saya dengan 3 paper teratas 5 tahun terakhir. Tunjukkan 3 celah yang **belum** disentuh dan cocok dengan data saya.”*

(4) Kesimpulan → (5) Abstrak → (6) Pendahuluan & Literature Review

Kita tulis **terbalik**. Kenapa? Karena kamu sudah pegang angka.

- **(4) Kesimpulan dulu (pendek & tajam):**
 - 3–5 butir bernomor.
 - Tiap butir ada **angka inti + dampak**.
Contoh: “Model A menurunkan waktu inferensi **56%** (49s→21.5s), cocok untuk perangkat low-end.”
- **(5) Abstrak setelahnya (150–250 kata):**
 - **Gap → Tujuan → Metode → Angka kunci → Dampak**.
 - Hindari kata umum seperti “*signifikan*” tanpa angka.
- **(6) Pendahuluan & literature review terakhir:**
 - Kamu sudah tahu **apa yang baru**; tinggal **jahit** ke literatur.
 - Bandingkan **apple-to-apple** (dataset, metrik, protokol).

- **Prompt menulis cepat:**

“Susun kesimpulan 4 butir dari tabel ini. Lalu ubah menjadi abstrak 180 kata. Setelah itu, buat 1 paragraf pendahuluan yang menekankan gap & kontribusi.”

(7) Judul Terakhir yang Meng-highlight Novelty

Judul adalah **pukulan akhir**. Bikin informatif, ringkas, dan *searchable*.

- **Rumus judul yang aman:**

[Topik/Objek] — [Metode/Pendekatan] — [Angka/Dampak] — [Domain]

Contoh: “Klasifikasi Citra Medis dengan *Lightweight CNN*: **+18.5 pts F1**, Latensi Turun **56%** pada Perangkat Low-End.”

- **Do & Don't:**

- **Do:** pakai kata kuat (*Akselerasi, Inovatif, Baru, Terobosan*) **jika** benar-benar didukung angka.
- **Don't:** klaim muluk tanpa data; hindari *clickbait* ilmiah.

- **Prompt generator judul:**

“Berikan 7 opsi judul yang menonjolkan (i) dua angka unggulan saya, (ii) metode, (iii) domain Indonesia; urutkan dari paling teknis ke paling populer.”

Ringkasannya (tempel di meja belajar)

- **Data dulu.** Tanpa data, jalan di tempat.
- **Tulis dari konkret → abstrak:** Hasil nyata dulu, teori belakangan.
- **Punya ≥ 3 novelty** yang **terukur**, bukan sekadar klaim.
- **Kesimpulan → Abstrak → Pendahuluan** biar hemat revisi.
- **Judul terakhir**, agar akurat dan meng-*highlight novelty*.

Mantra: Data memimpin, AI mempercepat. Kamu tetap peneliti utamanya.

Bagian I — Menyiapkan Bahan: Eksperimen, Observasi, dan Data

Merancang Eksperimen & Observasi yang “Ramah AI”

Kalau kamu ingin tulisan ilmiahmu **kuat dan anti-halusinasi**, maka semua harus dimulai dari **data yang rapi dan bisa dilacak**.

Ingat: *AI itu cuma asisten pintar, bukan cenayang.*

Kalau datamu berantakan, AI hanya akan menebak — dan hasilnya bisa ngawur.

Jadi, sebelum kamu mulai menulis, yuk kita bahas cara bikin eksperimen atau observasi yang *ramah AI* — alias data-friendly, bisa dibaca, dan gampang dipakai ulang.

1. Tentukan Variabel, Kontrol, dan Protokol

Setiap eksperimen selalu punya **tiga elemen penting**:

- **Variabel** → hal yang kamu ubah (misalnya: suhu, durasi, atau model yang digunakan).
- **Kontrol** → kondisi pembanding yang tetap.
- **Protokol** → langkah-langkah eksperimen, seperti resep masakan.

Contoh sederhana:

Kamu ingin tahu seberapa cepat *model AI* memproses gambar.

- Variabel: jenis model (ResNet, EfficientNet, MobileNet).
- Kontrol: dataset yang sama.
- Protokol: waktu inferensi diukur tiga kali dan dirata-ratakan.

Tips: tulis protokol mu seperti *cheat sheet*, supaya siapapun (termasuk kamu di masa depan) bisa mengulanginya tanpa bingung.

2. Catat Data Secara Digital dan Terstruktur

AI tidak bisa membaca tulisan tangan yang miring-miring di buku catatan. Maka, biasakan semua data disimpan dalam **format digital standar**, seperti:

- **.CSV** (mudah dibuka di Excel atau Python Pandas),
- **.Parquet** (lebih efisien untuk data besar).

Gunakan kolom dan nama yang **konsisten**, misalnya:

```
timestamp, model_name, accuracy, latency_ms
```

Pro-tip: Simpan juga *log* dari alat, hasil percobaan, atau output dari program. File `.log`, `.txt`, atau `.json` bisa sangat membantu ketika kamu ingin melacak asal hasil tertentu.

3. Gunakan Skema Penamaan Berkas yang Jelas

Banyak mahasiswa kehilangan data bukan karena rusak, tapi karena **bingung nama file sendiri**.

Daripada `data_final_fix_v2_beneran.csv`, lebih baik pakai format seperti:

```
YYYYMMDD_deskripsi_singkat.csv  
2025-03-10_uji_resnet_cifar10.csv
```

Kalau eksperimenmu banyak, buat **struktur folder seperti ini**:

```
/data_raw/  
/data_clean/  
/analysis_scripts/  
/results/  
/figures/
```

Dengan begitu, kamu bisa bilang ke AI:

“Analisis file di folder `/data_clean/` dan buat ringkasan korelasi antar variabel.”

4. Buat *Metadata Sheet*

Metadata itu semacam “KTP”-nya dataset kamu. Isinya menjelaskan **kapan, di mana, dan bagaimana data dikumpulkan**. Tanpa metadata, data kamu seperti barang hilang di dunia digital — tidak punya asal-usul.

Isi minimal metadata:

Kolom	Keterangan
data_source	Dari mana data berasal
date_collected	Tanggal dan waktu pengambilan
device_or_tool	Alat atau software yang digunakan
parameter_version	Versi model atau algoritma
collector	Siapa yang mengumpulkan data

Kenapa penting?

Karena saat AI membantu kamu menulis hasil atau analisis, metadata ini bisa digunakan untuk **mencegah kesalahan tafsir**. Misalnya, AI bisa tahu data lama atau baru, dan menyesuaikan konteksnya.

5. Kotak Prompt: Minta AI Memeriksa Kelengkapan Data Provenance

Nah, setelah semua siap, kamu bisa minta AI bantu mengecek apakah data kamu sudah lengkap dan siap dipakai.

Prompt contoh:

“Cek file CSV ini. Apakah sudah punya metadata yang cukup untuk diulang oleh peneliti lain?
Sebutkan apa yang kurang dan beri saran perbaikan.”

AI akan menilai hal-hal seperti:

- Apakah kolom *timestamp* ada?
- Apakah alat dan versi tercatat?
- Apakah hasil bisa dilacak kembali ke sumber?

Ini disebut **audit data provenance**, yaitu kemampuan melacak *asal-usul data*.

Kalau setiap angka bisa ditelusuri kembali ke alat, waktu, dan kondisi eksperimen, maka tulisanmu **bebas halusinasi ilmiah**.

6. Tantangan untuk Mahasiswa

Coba lakukan eksperimen mini:

1. Catat suhu ruangan atau waktu respon laptop kamu setiap 10 menit selama 1 jam.
2. Simpan dalam format **.csv**.
3. Buat metadata sheet sederhana.
4. Minta AI menganalisis pola dari data itu.

Kalau AI bisa menjelaskan hasilmu tanpa error, berarti datamu sudah **ramah AI!**

Kesimpulan Singkat

- **Eksperimen yang baik = data yang bisa diulang.**
- **AI hanya secerdas data yang kamu berikan.**
- Gunakan format digital, penamaan rapi, dan metadata lengkap.
- Biarkan AI jadi **asisten validasi data**, bukan tukang karang angka.

Mantra Bab Ini:

“Tanpa data yang rapi, AI cuma bisa menebak.”

Maka, jadilah peneliti yang bukan hanya pintar menulis, tapi juga **rapi, transparan, dan siap diaudit.**

Kebersihan Data & Dokumentasi

Kamu mau naskah ilmiah yang **kuat**? Mulailah dari data yang **bersih**, **rapi**, dan **terdokumentasi**. Tanpa itu, analisismu akan goyah. AI pun ikut bingung. Ingat: **data yang bersih = argumen yang tajam**.

1) Validasi, *Outlier*, dan *Missingness* (tiga serangkai wajib)

a) **Validasi** = cek “sehat”-nya data.

Tujuan: pastikan format benar, nilai masuk akal, tidak ada yang aneh.

- **Checklist validasi cepat:**
 - Tipe kolom konsisten? (angka ya angka, tanggal ya tanggal)
 - Rentang nilai masuk akal? (suhu 10–45°C, bukan 450°C)
 - Nilai *duplicate*? Jika ya, apakah wajar?
 - Satuan konsisten? (cm vs m; detik vs ms)
- **Contoh kalimat validasi:**

“Kolom `latency_ms` berada di rentang **[18, 120]**, tidak ada nilai negatif.”

b) **Outlier** = nilai yang “menyolok”. Bisa error, bisa juga “emas”.

Jangan langsung buang. **Amati dulu**.

- **Langkah simpel:**
 - Plot **boxplot** atau **histogram**.
 - Cek sumber: salah input? sensor error? percobaan sah tapi ekstrem?
 - Tulis keputusan: **hapus** / **perbaiki** / **simpan** (dengan alasan).
- **Contoh keputusan:**

“Nilai `latency_ms=2500` diduga *logging error*. Dihapus dari analisis utama, disimpan di *appendix*.”

c) **Missingness** = data hilang (*NaN/null*).

Ini wajar. Yang penting **terkendali**.

- **Strategi cepat:**
 - Jika <5% hilang: **hapus baris** (jika aman).
 - Jika penting: **imputasi** (rata-rata/median/kategori “unknown”).
 - Catat cara kamu menangani *missingness*. Transparansi itu penting.
- **Contoh catatan:**

“`temperature` hilang **3.1%**. Di-*impute* dengan median per sesi.”

Inti: Bersihkan **tanpa merusak cerita**. Semua keputusan **ditulis**, bukan sekadar “katanya”.

2) *Versioning* dengan *Git* / *DVC* (biar jejakmu rapi & terulang)

Kamu perlu **jejak perubahan**. Supaya data & skrip bisa **diulang** dan **dicek**.

- **Apa itu *Git*?**
Sistem *version control* untuk **kode & catatan**. Kamu bisa *commit*, *branch*, dan kembali ke versi lama.
- **Apa itu *DVC (Data Version Control)*?**
Mirip *Git*, tapi fokus pada **data & model** yang ukurannya besar. DVC menyimpan **pointer** (bukan menaruh data besar di repo), rapi dan hemat.
- **Workflow singkat (aman untuk pemula):**
 - Simpan kode/skrip dan catatan di *Git*.
 - Simpan data mentah di folder `/data_raw/`, data bersih di `/data_clean/`.
 - Lacak data dengan *DVC* (supaya bisa kembali ke versi tertentu).
 - Tulis ringkas di `README.md`: dari **raw** → **clean** → **hasil**.
- **Penamaan *commit* yang jelas (hindari “final_fix_bener2”):**
 - `feat: tambah pembersihan missingness suhu`
 - `fix: hapus outlier latency > 2000ms`
 - `data: update /data_clean/latency_2025-03-10.parquet`

Benefit:

Reviewer tanya “angka ini dari mana?” → kamu **bisa telusuri**. Timmu butuh ulang eksperimen? → tinggal **checkout** versi.

3) Template Lab Notebook Digital (pendek, padat, enak dibaca)

Lab notebook digital = “buku harian eksperimen” versi modern. Bisa di Google Docs, Notion, Obsidian, atau Markdown.

Gunakan template 7-baris ini setiap kali eksperimen:

1. **Tujuan singkat:** apa yang dites hari ini?
2. **Data & sumber:** file apa? folder mana? (link/relatif path)
3. **Alat & versi:** *software/library* + versi (mis. *Python 3.11*, *pandas 2.2*)
4. **Protokol ringkas:** langkah 1–4, cukup satu baris tiap langkah
5. **Keputusan pembersihan:** validasi, *outlier*, *missingness* (plus alasan)
6. **Hasil cepat:** 2–3 angka kunci (mean/median/delta)
7. **Catatan lanjutan:** ide besok, anomali, *to-do*

Contoh isi (mini):

- Tujuan: cek pengaruh *preprocessing* X pada latensi.
- Sumber: `/data_raw/log_2025-03-10.csv` → `/data_clean/latency_clean.parquet`
- Alat: *Python 3.11*, *pandas 2.2*, *numpy 2.0*
- Protokol: hapus duplikat; *impute* median untuk `temperature`; simpan Parquet.

- Pembersihan: hapus *outlier* `latency_ms > 2000` (sensor freeze).
- Hasil cepat: rata-rata **84.2 ms**, median **80.5 ms**, Δ **-12%** vs baseline.
- Lanjutan: cek *subgroup* device lama vs baru.

4) Checklist Siap-Analisis (tempel di dinding)

- **Validasi lulus:** tipe kolom, rentang nilai, satuan konsisten.
- **Outlier:** diidentifikasi → diputuskan (hapus/perbaiki/simpan).
- **Missingness:** dihitung %, ditangani, **dicatat metodenya**.
- **Struktur folder rapi:** `/data_raw/`, `/data_clean/`, `/figures/`, `/analysis_scripts/`.
- **Versioning:** *Git* untuk kode & catatan; *DVC* untuk data besar.
- **Lab notebook:** diisi lengkap 7-baris.
- **README:** jelaskan alur *raw* → *clean* → *result*.

Yang penting: semua keputusan **tertulis**. Tanpa dokumentasi, kebersihan data tidak terlihat.

5) Tantangan 20 Menit (serius tapi fun)

1. Ambil satu file `.csv` milikmu.
2. Hitung *missingness* per kolom (berapa %).
3. Deteksi *outlier* sederhana (pakai aturan **$Q1-1.5 \times IQR$** atau nilai ekstrem manual).
4. Tulis keputusan pembersihan **1 paragraf** di *lab notebook*.
5. Simpan hasil bersih sebagai `.parquet`.
6. Buat *commit Git* dengan pesan jelas.

Selesai? Kamu sudah membuat dataset **siap-analisis**, **ramah AI**, dan **siap diaudit**. Mantap!

Ringkasnya

- **Validasi–outlier–missingness** adalah fondasi kebersihan data.
- **Versioning** dengan *Git/DVC* menjaga jejak dan replikasi.
- **Lab notebook digital + checklist siap-analisis** = kerja ilmiah yang rapi.
- Dengan kebersihan dan dokumentasi, kamu tinggal bilang ke AI:
“Ini data bersihku. Tolong analisis dengan tenang.” ✨

Mantra bab ini: Bersihkan. Dokumentasikan. Versikan. Ulangi.

Etika & Kepatuhan: Pondasi Kepercayaan Ilmiah

Meneliti itu bukan cuma *pintar ngitung*. Ini juga soal **tanggung jawab**. Kalau etika dilanggar, **hasil secanggih apa pun bisa ditolak**. Ingat: **reputasi ilmiah dibangun oleh kejujuran dan transparansi**.

1) Izin Etik: “Surat Jalan” Penelitianmu

Kalau penelitian menyentuh manusia, hewan, atau data sensitif, **minta izin etik** dulu. Tanpa itu, kamu seperti balap liar—cepat, tapi **berbahaya**.

Kapan wajib izin etik?

- Survei/eksperimen ke **manusia** (mahasiswa, pasien, responden).
- Data berisi info **identitas pribadi**.
- Intervensi ke **hewan uji**.
- Observasi yang bisa menimbulkan **risiko** (fisik/psikologis).

Yang harus kamu siapkan:

- **Tujuan jelas** dan **manfaat** untuk peserta.
- **Risiko** (kalau ada) + **cara mitigasi**.
- **Informed consent**: peserta tahu, paham, dan **setuju**.
- Mekanisme **keluar kapan saja** tanpa sanksi.
- Rencana **penyimpanan & penghapusan data**.

Mantra: No consent, no data. *Tanpa persetujuan, jangan sentuh datanya.*

2) Privasi Data: Lindungi Identitas, Amankan Akses

Privasi itu *non-negotiable*. Kamu **wajib** menjaga data peserta.

Prinsip sederhana:

- **Minimization**: ambil data **seperlunya aja**.
- **Pseudonymization**: ganti **nama** → **ID_acak**.
- **Anonymization**: hilangkan jejak identitas (alamat, NIK, email).
- **Access control**: data mentah hanya untuk tim yang **perlu**.

Praktik aman:

- Simpan di **folder terproteksi** (akses terbatas).
- Jangan unggah data sensitif ke *cloud* publik tanpa izin.
- Pisahkan **/data_identitas/** dan **/data_analisis/**.
- Gunakan format **CSV/Parquet** yang **bersih** dari kolom rahasia.

Contoh kalimat aman di paper:

“Data telah **dipseudonimkan**. Identitas peserta **tidak** muncul pada publikasi. *Consent* tertulis disimpan oleh tim riset.”

Catatan: AI boleh bantu bersihkan data, tapi **jangan** kirim **identitas asli** ke *model publik*. *Masking* dulu, selalu.

3) **Disclosure Alat & Versi Model: Transparansi yang Bisa Diulang**

Reviewer akan tanya: “Kamu pakai **alat apa? Versinya** berapa?” Jawabanmu harus **tepat**. Ini kunci **reproducibility**.

Apa yang perlu dicatat dan ditulis di paper?

- **Perangkat keras:** CPU/GPU/RAM (contoh: *RTX 3060, 12GB*).
- **Perangkat lunak:** OS, *Python/R/Matlab, library + versi*.
- **Pipeline analisis:** langkah ringkas dari **raw** → **clean** → **hasil**.
- **Model/alat AI:** nama, versi, *hyperparameter, random seed*.
- **Tanggal** pemrosesan data (kadang hasil beda karena *update*).

Format *disclosure* yang rapih (contoh singkat):

- *Environment:* **Ubuntu 22.04, Python 3.11**
- *Libraries:* **pandas 2.2, numpy 2.0, scikit-learn 1.5**
- *Model:* **XGBoost 2.0, seed=42, learning_rate=0.1, max_depth=6**
- *Hardware:* **Intel i7-12700, 32GB RAM, RTX 3060 (12GB)**

Kenapa penting?

Supaya peneliti lain bisa **mengulang** dan **menguji** klaim mu. Tanpa *disclosure*, naskahmu **sulit dipercaya**.

4) **Mini-Checklist Etika & Kepatuhan (tempel di meja!)**

- **Izin etik** sudah disetujui (atau dinyatakan **tidak perlu** oleh komite).
- **Informed consent** jelas: tujuan, risiko, hak peserta.
- **Privasi terjaga:** *minimization, pseudonymization/anonymization*.
- **Akses data dibatasi** (izin & kontrol).
- **Disclosure lengkap:** alat, versi, *hyperparameter, seed*, tanggal proses.
- **Dokumentasi rapi:** folder, README, *lab notebook*.

5) **Contoh *Template* “Blok Etika” (siap tempel ke metodologi)**

Etika & Kepatuhan — Penelitian ini telah mendapat **persetujuan etik** dari [Nama Komite, No. surat, Tanggal]. Semua peserta menandatangani **informed consent**. Data dikumpulkan dengan prinsip **minimization** dan disimpan dalam bentuk **pseudonim**; kolom identitas disimpan terpisah dan **tidak** digunakan dalam analisis. Analisis dilakukan pada **Ubuntu 22.04, Python 3.11, pandas 2.2, scikit-learn 1.5**. Model yang digunakan adalah **XGBoost 2.0** (*seed=42*).

Spesifikasi perangkat keras: **Intel i7-12700, 32GB RAM, RTX 3060 (12GB)**.
Pipeline analisis dapat direplikasi dari skrip pada folder `/analysis_scripts/`.

6) Tantangan 15 Menit (seru tapi serius)

1. Tulis **1 paragraf** *informed consent* untuk survei sederhana di kampus.
2. Buat **tabel mapping**: **nama** → **ID_acak**. Simpan terpisah.
3. Buat **snippet disclosure alat & versi** untuk proyekmu saat ini.
4. Isi **checklist** di atas. Yang kosong → jadi *to-do* minggu ini.

Goal: kamu siap audit kapan saja. **Data aman, proses transparan, klaim kuat.**

Ringkasnya

- **Izin etik** itu **wajib** kalau menyentuh manusia/hewan/data sensitif.
- **Privasi** bukan fitur tambahan, tapi **syarat utama**.
- **Disclosure alat & versi model** membuat risetmu **bisa diulang**.
- Etika yang kuat = **naskah dipercaya + proses lancar**.

Mantra bab ini: Amankan manusia. Amankan data. Jelaskan alat. Setelah itu, baru kita bicara **angka** dan **novelty**.

Bagian II — Dari Data ke Metodologi & Hasil

Analisis Data Inti (EDA Cepat & Tajam)

Apa itu EDA?

EDA (Exploratory Data Analysis) adalah langkah “ngobrol dengan data”. Kita **lihat pola**, **cari anomali**, dan **tulis angka kunci**. Tujuan: **mengerti data dulu**, baru bikin klaim.

Ingat: angka dulu, tafsir belakangan. *Plot* yang rapi, kalimat pendek, metrik jelas.

1) Distribusi: “Data kamu menyebar ke mana?”

Kita ingin tahu bentuk sebaran nilai. Sempel tapi penting.

- **Langkah cepat:**
 - Lihat **rata-rata**, **median**, **IQR** (rentang tengah).
 - Pakai *histogram* atau *boxplot* (hindari *chartjunk*).
 - Catat **skew** (miring ke kanan/kiri) dan **puncak** (bimodal?).
- **Kalimat hasil (format 1–2 baris):**
“Akurasi median **81.2%** (IQR **78.5–84.0**). Distribusi **sedikit miring kanan**, ada ekor panjang.”
- **Mini-checklist distribusi:**
 - Ada *outlier* ekstrem?
 - Skala dan satuan konsisten?
 - Perlu transformasi (*log*, *min-max*)?

Kotak Prompt — Distribusi

“Ringkas distribusi kolom **accuracy**. Sebutkan median, IQR, kemiringan, dan 1 saran transformasi bila perlu.”

2) Tren Waktu: “Naik, turun, atau musiman?”

Untuk data berurut waktu, lihat **perubahan** dan **pola periodik**.

- **Langkah cepat:**
 - Buat *line plot* sederhana.
 - Hitung **rata-rata bergerak** (*rolling mean*) 7/30 poin.
 - Cari **perubahan mendadak** (*change point*).
- **Kalimat hasil:**
“Latensi turun **12%** dalam 4 minggu (98.4 → **86.5 ms**). Ada lonjakan **+35 ms** pada 2025-03-10 (dugaan *logging*).”
- **Sinyal musiman?**
Gunakan *dekomposisi sederhana* (tren + musiman + sisa).

Kotak Prompt — Tren Waktu

*“Dari **latency_ms** per hari, deteksi 2 tren utama dan 1 tanggal anomali. Beri penjelasan 1 baris per poin.”*

3) Korelasi: “Siapa yang jalan bareng?”

Korelasi membantu menebak hubungan awal (bukan sebab-akibat ya!).

- **Langkah cepat:**
 - Hitung **korelasi Pearson** (dua variabel numerik).
 - Untuk kategorikal vs numerik: bandingkan **rata-rata per kategori**.
 - Lihat *scatter plot* untuk pola non-linear.
- **Kalimat hasil:**

“Korelasi **ukuran_model** vs **latency_ms** **r=0.78** (kuat). **batch_size** besar menurunkan latensi **-14%** rata-rata.”
- **Hati-hati:** korelasi **bukan** kausalitas. Hindari klaim berlebihan.

Kotak Prompt — Korelasi

“Hitung korelasi 5 pasang variabel teratas. Buat 3 kalimat: 2 korelasi kuat ($r > |0.6|$) + 1 korelasi lemah yang mengejutkan.”

4) Segmentasi (*Subgroup Analysis*): “Adil di semua kelompok?”

Cek performa di **kelompok berbeda** (device lama vs baru, daerah A vs B, dsb).

- **Langkah cepat:**
 - Bagi data ke **subgroup** yang relevan.
 - Laporkan **metrik per subgroup** (mean/median, error).
 - Hitung **gap** antar subgroup (dalam %).
- **Kalimat hasil:**

“Akurasi pada *device low-end* **76.3%**, *high-end* **84.7%** ($\Delta=+8.4$ pts). Perlu optimasi untuk low-end.”
- **Tujuan:** temukan **ketidakadilan performa** sedini mungkin.

Kotak Prompt — Segmentasi

*“Bandingkan akurasi per **device_tier**. Laporkan median tiap tier dan Δ antar-tier. Sarankan 1 eksperimen untuk mengurangi gap.”*

5) Anomali: “Nilai nakal yang perlu perhatian”

Anomali bisa **sampah**, bisa juga **penemuan**.

- **Cara cepat mendeteksi:**
 - Aturan **IQR**: nilai $< Q1 - 1.5 \times IQR$ atau $> Q3 + 1.5 \times IQR$.
 - **Z-score** > 3 atau < -3 untuk data normal.
 - Di waktu: lonjakan lokal dibanding *rolling median*.
- **Kalimat hasil:**

“Terdeteksi **7 anomali** ($Z > |3|$). 5 terkait *sensor freeze*, 2 terjadi saat *deployment* versi baru.”

Kotak Prompt — Anomali

“Tandai 5 nilai paling anomali di *Table X*. Jelaskan kemungkinan penyebab (*alat, data masuk, event*). Rekomendasikan tindakan cepat.”

6) Format Paragraf Hasil (3 kalimat aja)

Gunakan pola **angka** → **perbandingan** → **makna awal**.

- **Template:**
 1. **Angka inti:** “Median akurasi **81.2%** (IQR **78.5–84.0**).”
 2. **Bandingkan:** “Model A **+18.5 pts** vs baseline.”
 3. **Makna awal:** “Indikasi fitur X efektif; butuh cek stabilitas di low-end.”

Tebalkan angka penting. Hindari kalimat umum seperti “*cukup baik*”.

7) Sprint EDA 30 Menit (praktikal untuk mahasiswa)

Menit 0–10: Distribusi → median, IQR, 1 plot.

Menit 10–18: Tren waktu → *rolling mean*, cari 1 *change point*.

Menit 18–24: Korelasi → 3 pasangan terkuat + 1 *scatter*.

Menit 24–30: Segmentasi → 2 subgroup + $\Delta\%$; catat 1 anomali.

Output akhir: 1 tabel ringkas + 3 kalimat hasil. Selesai.

8) Kotak Prompt (siap tempel di proyek)

1. Ringkas EDA Umum

*“Baca **Tabel/Gambar X**. Ringkas: (a) median & IQR tiap metrik, (b) 2 tren waktu, (c) 2 korelasi teratas, (d) 1 gap subgroup, (e) 1 anomali penting.”*

2. Deteksi Pola/Anomali Paling Signifikan

*“Deteksi pola/anomali paling signifikan dari **Tabel/Gambar X**. Tulis 5 poin: metrik, nilai, konteks, dugaan penyebab, aksi lanjut.”*

3. Konversi ke Kalimat Ilmiah

“Ubah temuan EDA ini menjadi 3 kalimat hasil: angka → perbandingan → makna awal. Tebalkan angka kunci.”

Ringkasnya

- **Distribusi** untuk bentuk data.
- **Tren waktu** untuk arah dan kejutan.
- **Korelasi** untuk pasangan yang jalan bareng.
- **Segmentasi** untuk keadilan performa.
- **Anomali** untuk temuan atau peringatan.

Mantra EDA: lihat → ukur → tulis singkat.

AI boleh bantu, tapi **kamu** yang memutuskan makna.

Menulis Metode yang Reproducible (Dibantu AI)

Metode itu seperti **resep masakan ilmiah**. Kalau kamu menulisnya asal-asalan, orang lain **nggak akan bisa masak ulang hasilmu**. Padahal, di dunia riset, **kehebatan bukan cuma hasilnya**, tapi juga **seberapa mudah orang lain bisa mengulangnya**.

Prinsip utama: *If it can't be reproduced, it didn't really happen.*

Jelaskan Alat, Parameter, dan Algoritma dengan Jelas

Bayangkan kamu membuat eksperimen yang keren, tapi lupa nulis alatnya apa. Reviewer pun langsung bingung: *"Ini pakai apa ya? versi berapa?"*

Hal yang wajib dicantumkan:

- **Perangkat keras (hardware):** CPU, GPU, RAM, atau alat lab yang dipakai.

Contoh: *"Eksperimen dijalankan di RTX 3060 (12GB), RAM 32GB, CPU i7-12700."*

- **Perangkat lunak (software):** versi sistem operasi, bahasa pemrograman, *library*, dan tools.

Contoh: *"Python 3.11, TensorFlow 2.16, Ubuntu 22.04, pandas 2.2."*

- **Parameter dan konfigurasi:** tulis yang kritikal saja (bukan semua).

Contoh: *learning_rate = 0.001, batch_size = 32, epochs = 50.*

Tips:

- Simpan semua di file `config.yaml` atau `params.json`.
- Gunakan *Git versioning* supaya tiap perubahan bisa dilacak.
- Saat menulis paper, buat tabel kecil seperti ini:

Komponen	Nilai/Versi	Catatan
Model	XGBoost 2.0	seed=42
Optimizer	Adam	LR=0.001
Framework	TensorFlow 2.16	Ubuntu 22.04
GPU	RTX 3060	12GB VRAM

Bonus: kalau pakai *AI tools* (seperti ChatGPT, Copilot, atau Claude), tulis juga versinya dan prompt penting yang kamu gunakan. Itu bagian dari *reproducibility audit*.

Ablation Plan: Uji Coba Bongkar Pasang Fitur

Ablation study itu seperti **eksperimen iseng tapi cerdas**. Tujuannya untuk tahu: **fitur mana yang paling berpengaruh?**

Cara sederhana:

1. Jalankan model lengkap dulu (*baseline full feature*).
2. Hapus satu komponen (misal satu fitur, atau satu lapisan).
3. Catat perubahan hasilnya.
4. Ulangi untuk fitur lain.

Contoh:

Model	Komponen Dihapus	Akurasi (%)
A (Full)	—	91.2
B	tanpa “Text Augmentation”	88.5
C	tanpa “Normalization”	83.4

Kesimpulan: “Normalization meningkatkan akurasi **+7.8 pts.**”
Ini data, bukan opini. Reviewer suka hal seperti ini!

Catatan:

- *Ablation* juga bisa dipakai untuk bandingkan arsitektur model.
- Gunakan AI untuk bantu meringkas hasil tabel jadi kalimat singkat:

“Fitur Normalization memberi kontribusi terbesar (+7.8 pts), sedangkan Augmentation memberi efek sedang (+2.7 pts).”

Sensitivity Analysis: Seberapa “Galau” Model Kamu?

Kadang model bekerja baik, tapi cuma di kondisi tertentu. Nah, *sensitivity analysis* membantu kamu tahu: **berapa jauh parameter boleh berubah tanpa bikin hasil kacau.**

Langkah cepat:

1. Pilih parameter penting (misal *learning_rate*, *batch_size*, *dropout*).
2. Ubah sedikit nilainya (+10%, -10%).
3. Catat efeknya pada metrik utama (akurasi, F1, RMSE, dst).

Parameter	Nilai	Akurasi (%)
<i>learning_rate</i> = 0.001		91.2
<i>learning_rate</i> = 0.0005		90.7
<i>learning_rate</i> = 0.01		78.5

4. Buat grafik sederhana. Kesimpulan: “Model sensitif terhadap *learning_rate* > 0.005. Direkomendasikan tetap di bawah nilai tersebut.”

Tujuannya:

- Mengetahui **parameter paling kritis**.
- Meningkatkan **stabilitas** eksperimen.
- Memberikan **bukti empiris** bahwa hasil bukan kebetulan.

Kotak Prompt — Minta AI Bantu Menulis Bagian Metode

Kamu bisa hemat waktu dengan meminta AI menulis bagian metode dari daftar alat/parameter. Tapi ingat: **AI bantu menulis, bukan menggantikan data!**

Contoh Prompt:

“Buat narasi metode penelitian dari daftar berikut:
Model: Random Forest (*n_estimators*=200, *max_depth*=10)
Dataset: 12.000 data bersih
Framework: scikit-learn 1.5
Hardware: Intel i7, 16GB RAM, Windows 11
Tambahkan catatan singkat soal ablation dan sensitivity.”

Output AI yang ideal:

“Penelitian ini menggunakan algoritma *Random Forest* dengan 200 pohon dan kedalaman maksimum 10. Model diimplementasikan menggunakan *scikit-learn 1.5* di atas lingkungan *Windows 11* dengan prosesor Intel i7 dan RAM 16GB. Untuk memastikan kestabilan hasil,

dilakukan uji *sensitivity analysis* terhadap parameter `max_depth` dan `n_estimators`, serta *ablation study* untuk mengevaluasi kontribusi tiap fitur.”

Kalau kamu lihat outputnya rapi, tinggal **verifikasi angka dan sumbernya**.

Tantangan 20 Menit (buat mahasiswa yang berani!)

1. Ambil proyek analisis datamu (boleh Tugas Akhir, boleh latihan kelas).
2. Buat tabel alat + parameter seperti di atas.
3. Buat *ablation plan* (3 skenario minimal).
4. Ubah 1 parameter → catat perubahan hasil (*sensitivity test*).
5. Tulis paragraf metode 5 kalimat (boleh pakai bantuan AI).

Goal: menghasilkan “metode yang bisa diulang” oleh temanmu — tanpa tanya-tanya kamu lagi.

Ringkasan

- Metode yang *reproducible* = riset yang dipercaya.
- Jelaskan **alat, versi, dan parameter penting**.
- Gunakan **ablation plan** untuk tahu bagian paling berpengaruh.
- Gunakan **sensitivity analysis** untuk uji stabilitas.
- Gunakan AI dengan bijak untuk bantu menulis narasi teknis.

Mantra bab ini:

“Kalau hasilmu nggak bisa diulang, berarti belum selesai.”

Jadi, tulis metode seperti *resep nasi goreng ilmiah*:
jelas, lengkap, dan bisa dimasak siapa saja.

Menulis Hasil: Angka Dulu, Tafsir Sesudah

Kenapa “angka dulu”?

Karena **hasil riset = bukti**, bukan opini.

Kita tunjukkan angka **apa adanya** dulu. Baru **bandingkan**. Lalu **tafsir** singkat.

Urutannya sederhana: **angka** → **perbandingan** → **makna awal**.

Kalau terbalik, risiko *bias* dan *halusinasi* naik.

1) Tabel Minimalis, Jelas, Anti-Ribet

Tujuan: pembaca paham dalam 10 detik.

Prinsip:

- **Kolom ringkas.** Maks 5–6 kolom inti.
- **Satuan konsisten.** Jangan campur %, detik, dan ms di kolom yang sama.
- **Tebalkan angka terbaik.** Boleh miringkan *runner-up*.
- **Tambahkan catatan uji.** (*n*, dataset, protokol).

Contoh pola tabel (teks):

Tabel 1. Akurasi (%) pada Dataset X (*n*=1.200, 3-fold CV)

Model	Mean	SD	Waktu (s)
Baseline	82.1	1.3	45
Model-A	88.7	1.1	52
Model-B	90.4	1.0	58

Tips: sertakan *SD/SE* untuk menandai variasi. Itu bikin klaim kamu **lebih kredibel**.

2) Grafik Tanpa *Chartjunk*

Chartjunk = hiasan tak perlu yang **mengganggu pesan**.

Hilangkan: bayangan 3D, pola batik, emoji di sumbu, dan warna pelangi yang tidak bermakna.

Grafik yang sehat:

- **Label sumbu jelas + satuan.**
- **Judul informatif.** “Akurasi per Model di Dataset X (*n*=1.200)”.
- **Error bar** untuk *confidence interval* 95% atau *standard error*.
- **Legenda hemat.** Maks 3–4 seri.

Pilih grafik sesuai tujuan:

- Banding nilai satu waktu → **bar/column** dengan *error bar*.
- Tren waktu → **line**.
- Distribusi → **histogram/boxplot**.

- Hubungan dua variabel → **scatter** + garis *fit*.

Ingat: grafik itu alat komunikasi, bukan karya seni. **Sederhana menang.**

3) Struktur Paragraf Hasil: 3 Kalimat Emas

Template 3-langkah (copas lalu isi angkanya):

1. **Angka (apa yang terjadi).**
“Model-B mencapai **90,4%** ($\pm 1,0$) pada Dataset X, lebih tinggi dari Baseline **82,1%**.”
2. **Perbandingan (seberapa besar bedanya).**
“Kenaikan relatif = $(90,4 - 82,1) / 82,1 \approx 10,1\%$; waktu komputasi naik **+13 s**.”
3. **Makna awal (kenapa penting).**
“Artinya, Model-B **lebih akurat** dengan biaya waktu yang **masih wajar** untuk skala praktis.”

Contoh lengkap (ringkas, padat, tajam):

“Pada Dataset X, Model-B mencapai **90,4%** ($\pm 1,0$) dibanding Baseline **82,1%**.

Kenaikan relatif $\approx 10,1\%$, dengan overhead waktu **+29%** (45 s → 58 s).

Hasil ini menunjukkan **keuntungan akurasi signifikan** dengan biaya komputasi yang **masih dapat diterima** untuk *production* ringan.”

Kuncinya: angka spesifik dulu, perbandingan terukur kemudian, tafsir singkat terakhir.

4) Checklist “Angka Dulu, Tafsir Sesudah”

- [] Ada tabel/grafik utama sebelum narasi.
- [] Metrik dan satuan jelas (mis. akurasi %, RMSE, detik).
- [] Ada *error bar* / SD / CI.
- [] Ada kalimat **perbandingan kuantitatif** (Δ absolut/relatif).
- [] Ada **konteks uji** (dataset, n , protokol).
- [] Tafsir hanya **1–2 kalimat** di akhir paragraf.

5) Kotak Prompt — Narasi Hasil dari Tabel/Grafik Mentah

Gunakan AI sebagai **asisten narasi**, bukan pengganti data.

Prompt 1 — Paragraf 3 Kalimat Emas

“Gunakan Tabel berikut untuk menulis paragraf hasil 3 kalimat:

1. angka inti; 2) perbandingan Δ absolut & relatif; 3) makna awal.
Sertakan metrik & satuan. Hindari opini.
[Tempel Tabel/angka mentah di sini]”

Prompt 2 — Ringkas Grafik + *Error Bar*

“Dari grafik (mean $\pm$ SD) ini, tulis 2–3 kalimat:
sebutkan nilai terbaik, selisih dengan Baseline, dan apakah perbedaannya besar secara praktis.
[Tempel deskripsi angka mean/SD per model]”

Prompt 3 — Ubah Kalimat Kualitatif → Kuantitatif

“Ubah pernyataan kualitatif berikut menjadi kuantitatif dengan rumus $\Delta\% = ((A-B)/B) \times 100\%$.
Laporkan nilai absolut, $\Delta\%$, dan *confidence interval* jika ada.
[Tempel dua angka yang dibandingkan]”

Prompt 4 — Satu Paragraf per Gambar

“Tulis narasi hasil untuk Gambar X: mulai dengan angka inti, lalu bandingkan model,
akhiri dengan makna awal. Maks 70 kata.
[Tempel ringkasan angka dari Gambar]”

Pro tip: setelah AI menulis, **cek ulang angka & satuan**. AI boleh bantu kalimat, **data tetap raja**.

6) Mini-Misi 15 Menit (buat kamu yang penasaran)

1. Ambil satu tabel hasil terbaru.
2. Buat grafik **tanpa *chartjunk* + *error bar***.
3. Tulis **1 paragraf** dengan template 3 kalimat.
4. Jalankan **Prompt 1** di atas, bandingkan dengan tulisanmu.
5. Pilih versi terbaik → **tempel** ke bagian *Results*.

Tujuan: kamu bisa menyajikan hasil yang **tajam, terukur, dan jujur**.

Ringkasannya

- Mulai dari angka.
- Bandingkan dengan jelas (Δ absolut & $\Delta\%$).
- Tafsir singkat di akhir.
- Tabel minimalis dan grafik bersih (tanpa *chartjunk*).
- Gunakan *prompt* untuk mempercepat narasi, tapi verifikasi semua angka.

Mantra: Data bicara dulu, kita bicara belakangan.

Bagian III — Menemukan & Memastikan Novelty

Definisi Novelty yang Terukur

1) Apa itu *novelty* yang terukur?

Novelty artinya **ada hal baru** yang kamu bawa. Bukan perasaan. Bukan dugaan. **Harus bisa diukur dan bisa diuji ulang.**

Kuncinya:

- **Operasional:** jelas *apa* yang baru, **bagaimana** mengukurnya, **dengan metrik apa, di kondisi apa.**
- **Teruji:** hasilnya **stabil** saat diulang (*repeatability*) dan **konsisten** saat dikerjakan tim/alat lain (*reproducibility*).
- **Terverifikasi:** bisa **dibandingkan** dengan studi lain atau *baseline* yang kuat.

Ringkasnya: “**baru + terukur + bisa diulang = *novelty* valid.**”

2) Jenis *novelty* (pilih minimal 3 jalur)

Kamu bisa membuktikan *novelty* lewat **lima jalur** ini. **Ambil 3 atau lebih** agar paper kamu solid.

a) Metode (*method novelty*)

Ada teknik/algoritma/langkah baru, atau kombinasi baru.

- **Contoh klaim terukur:**
“**Error turun 18,7%** dengan **metode-X + *data augmentation* Y** dibanding baseline Z pada protokol A.”

b) Data/Domain

Ada **dataset baru**, domain baru, atau **kondisi uji** baru yang penting.

- **Contoh klaim terukur:**
“Model-B **tetap stabil** ($SD \leq 1,2$) pada **lingkungan tropis** yang belum diuji studi terdahulu.”

c) Performa

Membawa **angka lebih baik** pada metrik yang relevan.

- **Contoh klaim terukur:**
“Akurasi naik dari **82,1%** → **90,4%** ($\Delta_{rel} \approx 10,1\%$), **n=1.200**, 3-fold CV.”

d) Efisiensi

Lebih **cepat**, lebih **irit komputasi/energi/biaya**, atau **lebih kecil modelnya**.

- **Contoh klaim terukur:**
“Waktu inferensi **turun 42%** (58 s → 33 s) dengan **memori -35%**, akurasi **turun <0,5%**.”

e) Interpretasi

Memberi **pemahaman baru**: *why/how* sebuah fenomena terjadi.

- **Contoh klaim terukur:**
“Fitur ‘humidity-index’ menjelaskan **31% variansi** hasil ($R^2=0,31$), **$p<0,01$** → **mekanisme baru didukung data**.”

Catatan: semua contoh pakai **angka, metrik, dan kondisi uji**.
Tanpa itu, *novelty* jadi kabur.

3) Cara memastikan *novelty* benar-benar terukur

Pakai *checklist* ini saat menulis klaim:

- **[] Definisi operasional** jelas (metrik, satuan, protokol, dataset).
- ****[] Ada perbandingan** dengan baseline/standar (Δ absolut & $\Delta\%$).
- ****[] Ada ketidakpastian:** SD/SE/CI, *n*, *split* data.
- **[] Repeatability** lolos (uji ulang oleh kamu → hasil konsisten).
- **[] Reproducibility** siap (orang lain bisa replikasi dengan info yang kamu tulis).
- **[] Relevansi** ke masalah nyata (kenapa “baru” ini **penting**).

Bold rule: tanpa angka dan protokol, klaim *novelty* = opini.

4) Target ≥ 3 *novelty* yang bisa diuji/diulang

Agar paper **tahan banting**, targetkan **tiga atau lebih *novelty*** dari kombinasi lima jalur di atas.

Strategi cepat (contoh paket 3 *novelty*):

- **Performa:** “A **+10% akurasi** vs baseline pada Dataset X.”
- **Efisiensi:** “Waktu training **-35%** di *hardware* mid-range.”
- **Interpretasi:** “Fitur Z **menjelaskan 28% variansi**; valid di 2 *subset* domain.”

Atau paket lain:

- **Metode + Data/Domain + Performa,**
- **Metode + Efisiensi + Interpretasi,**
- **Data/Domain + Performa + Efisiensi.**

Format penyajian (ringkas & tembus pandang):

- **Novelty-1 (Performa)** — metrik, Δ , protokol, n , CI.
- **Novelty-2 (Efisiensi)** — waktu/energi/memori, Δ , perangkat uji.
- **Novelty-3 (Interpretasi)** — R^2 /effect size/p-value, variabel kunci.

5) Mini-misi 10 menit: *Novelty Scorecard*

Biar cepat, isi matriks kecil ini dari datamu sendiri.

Tabel *scorecard* (contoh isi):

Aspek	Metrik & Kondisi	Nilai	Δ vs Baseline	Uji Ulang
Performa	Akurasi (%), Dataset X, 3-fold, $n=1200$	90,4 $\pm$ 1,0	+10,1%	Lolos ($\pm 1\%$)
Efisiensi	Waktu (s), GPU mid-range	33	-42%	Lolos
Interpretasi	R^2 fitur Z, p-value	0,31 ($p < .01$)	-	Lolos

Goal: punya ≥ 3 baris dengan isi lengkap & terukur. Itu *novelty* inti paper kamu.

6) Do & Don't *novelty*

Do

- **Spesifik:** sebut metrik, angka, protokol, dataset.
- **Bandingkan** dengan baseline kuat (paper/top-3 metode).
- **Laporkan** ketidakpastian (CI/SD/SE) dan n .
- **Sertakan** batasan kecil $\rightarrow$ terlihat jujur dan ilmiah.

Don't

- Klaim “lebih baik” **tanpa angka**.
- Ganti baseline setiap tabel (tidak *fair*).
- Pindah metrik seenaknya (cherry-picking).
- Hanya “feels new” $\rightarrow$ **tanpa bukti**.

7) Kalimat siap pakai (isi angkanya)

- “Metode kami **meningkatkan akurasi** dari **B%** $\rightarrow$ **A%** ($\Delta_{rel} = ((A-B)/B) \times 100\%$) pada **Dataset X**, $n=...$, **3-fold**.”
- “Waktu inferensi **turun C%** ($D \rightarrow E$ detik) di **GPU/CPU** yang sama.”
- “Variabel Z **menjelaskan F% variansi** ($R^2=F/100$), $p<...$, menunjukkan **mekanisme baru** yang konsisten.”

8) Kotak refleksi (1 menit)

Tanya dirimu:

- **Apa yang benar-benar baru** dari kerjamu?
- **Di angka berapa** “barunya” terlihat?
- **Jika diulang**, apakah hasilnya **tetap muncul**?

Kalau tiga pertanyaan ini bisa kamu jawab **dengan angka**, kamu punya **novelty** yang terukur.

Kesimpulan mini:

Novelty yang kuat = (baru) + (terukur) + (bisa diulang).

Ambil ≥ 3 jalur novelty dari metode, data/domain, performa, efisiensi, interpretasi.

Tulis dengan angka, metrik, protokol, dan bandingkan secara adil.

Itu resep paper yang tajam, kredibel, dan susah dipatahkan.

Cek Novelty dengan AI (Anti-Halusinasi)

Kenapa perlu cek *novelty* pakai AI?

AI itu **cepat** dan **tajam**. Tapi AI juga bisa **halusinasi**. Maka kita pakai AI sebagai **detektor peluang**, bukan **penentu akhir**. Prinsipnya: **AI cari, kita verifikasi**. Kalau AI bilang “mirip paper X”, tugas kita: **buka sumbernya, cek DOI, cek angka**.

Rule emas: Tanpa bukti (DOI, halaman, tabel, angka), anggap klaim AI belum valid.

Strategi 5 Tahun Terakhir (biar relevan & *state-of-the-art*)

Targetkan literatur di **5 tahun terakhir**. Kenapa? Karena:

- Metodologi cepat berubah.
- Standar metrik ikut update.
- Dataset baru bermunculan.

Checklist cepet:

- [] Batasi pencarian: **tahun $\geq$ (tahun sekarang – 5)**.
- [] Prioritaskan jurnal/prosiding **terindeks** (*Scopus, Web of Science, PubMed, IEEE Xplore*).
- [] Simpan **3–10** kandidat teratas untuk dibedah.

Prompt siap pakai:

- “Cari 10 artikel jurnal/prosiding terindeks tentang [topik], tahun $\geq$ 2020, tampilkan judul, DOI, sumber, dan 1–2 kalimat temuan kunci.”
- “Kelompokkan artikel itu per jenis kontribusi: metode, data/domain, performa, efisiensi, interpretasi.”

Forward/Backward Citations (menyisir jaringan kutipan)

Bayangkan satu paper penting sebagai “node”. Kita telusuri dua arah:

- **Backward citations:** daftar referensi **yang dikutip** oleh paper itu.
Guna: melihat **akar teori** dan standar lama yang dipakai.
- **Forward citations:** daftar paper **yang mengutip** paper itu.
Guna: menangkap **perkembangan terbaru** dan *improvement* setelahnya.

Langkah ringkas:

1. Ambil 1–3 paper kunci (paling relevan atau paling banyak disitasi).
2. Baca **tabel/hasil** dan **metode** mereka. Catat **metrik, dataset, protokol**.
3. **Backward:** adakah teknik yang *under-explored*?

4. **Forward:** adakah paper baru yang **mengalahkan** angka mereka? Di **kondisi** apa?

Prompt siap pakai:

- “Dari paper [DOI/judul], daftar 5 rujukan paling penting (*backward*) beserta metrik kunci yang dipakai.”
- “Cari 5 paper terbaru (*forward*) yang mengutip [DOI/judul] dan laporkan apakah ada perbaikan metrik/perubahan protokol.”

Validasi Sumber: DOI, Penerbit, Indeks (Anti-Halu Mode ON)

Ini adalah **mode aman** agar AI tidak menipu kita dengan referensi palsu.

Tiga lapis verifikasi:

1. **DOI valid**
 - Harus ada **format DOI** yang bisa dilacak.
 - Cocokkan judul, penulis, tahun dengan halaman penerbit.
2. **Penerbit kredibel**
 - Cek halaman **publisher** (mis. *IEEE, Springer, Elsevier, ACM, Nature, Wiley, MDPI*—tetap kritis).
 - Baca **PDF asli**, bukan blog/SlideShare.
3. **Terindeks**
 - Ada di **Scopus/Web of Science/PubMed/IEEE Xplore** (sesuai bidang).
 - Kalau *preprint* (*arXiv*), beri label jelas: **belum peer-reviewed**.

Checklist anti-halu:

- [] DOI bisa dibuka & cocok judul/penulis.
- [] Link ke halaman penerbit resmi.
- [] Status indeks jelas (Scopus/WoS/... atau preprint).
- [] Angka & tabel **benar-benar ada** di paper (sebut **nomor tabel/gambar/halaman** saat mengutip).

Prompt siap pakai:

- “Verifikasi referensi berikut (judul, DOI, penerbit, terindeks apa). Melaporkan ketidaksesuaian dan tandai mana yang *preprint*.”

Bandingkan Temuan Kita vs 3 Paper Teratas (apples-to-apples)

Tujuan: pastikan *novelty* kita **benar-benar baru** pada **metrik & kondisi** yang sama.

Langkah praktis:

- 1. Pilih **3 paper teratas** (yang paling relevan & kredibel).
- 2. Ekstrak **dataset, metrik, protokol, hardware, ukuran sampel**.
- 3. Samakan **protokol uji** sedekat mungkin.
- 4. Taruh **angka kita** sejajar di tabel perbandingan.
- 5. Tulis **$\Delta\%$** dan **ketidakpastian (SD/CI)**.
- 6. Simpulkan **apa yang baru** dan **di kondisi apa**.

Template tabel ringkas:

Studi	Dataset/Domain	Metrik (protokol)	Nilai ($\pm$ CI)	Catatan Kunci
Paper-A (2023)	X (split resmi)	Akurasi, 5-fold, n=1200	88,2 $\pm$ 0,9	Baseline kuat
Paper-B (2022)	X (varian B)	F1, holdout 20%, n=900	0,842 $\pm$ 0.02	Lebih baik di F1
Paper-C (2024)	X (noisy)	RMSE, 3-run avg	0,214 $\pm$.01	Fokus regresi
Kita (2025)	X (protokol A)	Akurasi, 5-fold, n=1200	90,4 $\pm$ 1,0	+10,1% vs A

Kalimat siap tempel (isi angkanya):

- “Dibanding **Paper-A (2023)** pada **Dataset X** dan **protokol yang sama**, akurasi kami naik **+10,1% (90,4 $\pm$ 1,0 vs 88,2 $\pm$ 0,9)**.”
- “Keuntungan bertahan pada **3 subset domain**, **Δ rel 8–12%, $p < 0,01$** .”

Kotak Prompt — Paket Siap Pakai

Gunakan *prompt* ini apa adanya. Isi bagian dalam kurung siku.

1) Pemetaan cepat 5 tahun terakhir

“Ringkas 10 artikel/jurnal terindeks tentang [topik] ($\geq$ [tahun-5]). Untuk tiap artikel, tampilkan: judul, DOI, penerbit, indeks (Scopus/WoS/...), dataset, metrik, protokol, angka kunci, dan 1 kalimat kontribusi.”

2) Forward/Backward citations

“Dari paper [DOI/judul], tampilkan 5 backward citations terpenting (dengan DOI dan angka kunci), dan 5 forward citations terbaru ($\geq$ [tahun-3]) yang memperbaiki hasil atau mengubah protokol.”

3) Validasi sumber (anti-halu)

“Validasi daftar referensi ini: cek DOI valid, cocok judul/penulis/tahun di halaman penerbit, dan status indeks. Tandai yang preprint. Laporkan anomali.”

4) Perbandingan apples-to-apples

“Bandingkan temuan kami [deskripsi singkat + angka] dengan 3 paper teratas berikut ([DOI-1], [DOI-2], [DOI-3]). Buat tabel: dataset, metrik, protokol, nilai $\pm$ CI, $\Delta\%$ vs kami. Lalu tulis 3 kalimat simpulan novelty yang **terukur**.”

5) Deteksi celah (gap) untuk dorong *novelty*

“Dari 3 paper pembandingan, identifikasi 3 celah: (a) domain/lingkungan belum diuji, (b) metrik kurang relevan, (c) efisiensi/biaya belum diukur. Sarankan eksperimen singkat untuk menutup celah dengan target angka.”

Do & Don't saat cek *novelty* dengan AI

Do

- **Batasi tahun** → fokus 5 tahun terakhir.
- **Sebut DOI** dan **buka PDF** penerbit.
- **Samakan protokol uji** sebelum mengklaim “lebih baik”.
- **Cantumkan CI/SD/SE** dan **n**.

Don't

- Mengutip **tanpa DOI**.
- Mencampur metrik/protokol lalu mengklaim menang.
- Percaya ringkasan AI **tanpa cek PDF**.
- *Cherry-pick* satu angka dan abaikan batasan.

Ringkasan kilat

- **AI = akselerator cek *novelty***, bukan hakim akhir.
- Pakai **strategi 5 tahun terakhir + forward/backward citations** untuk peta bidang.
- **Validasi sumber** ketat: **DOI–penerbit–indeks**.
- **Bandingkan** temuan kita **vs 3 paper teratas** secara **apples-to-apples** (dataset, metrik, protokol).
- Tulis klaim **dengan angka** dan **ketidakpastian**.
- Hasilnya: *novelty* yang **terukur, kredibel, dan anti-halusinasi**.

Ayo praktik: ambil satu *result* utama dari datamu, jalankan **Kotak Prompt #4**, dan lihat apakah klaimmu **benar-benar baru**. Kalau belum, pakai **Kotak Prompt #5** untuk membuka celah *novelty* baru.

Jika Novelty Lemah: Iterasi Eksperimen Terarah

Kenapa harus iterasi?

Kadang hasil kita “oke”, tapi **belum beda**. Itu normal. Kuncinya: **jangan ganti semuanya**, cukup ubah yang **paling berpengaruh**. Tujuan kita jelas: **mengangkat sinyal *novelty*** dengan biaya dan waktu minimum.

Prinsip emas: kecilkan ruang tembak → bidik variabel kritis → naikkan efek → buktikan dengan angka.

Empat tuas untuk mengangkat *novelty*

1. Ubah variabel kritis

Contoh: suhu reaksi, *learning rate*, jumlah lapisan, konsentrasi, jenis sensor.

- **Mulai dari 1–2 variabel** yang paling sensitif.
- Pakai desain sederhana: **1-Factor-At-a-Time** dulu; jika perlu, lanjut *factorial design*.

2. Uji di domain baru

Pindah konteks sedikit, dampaknya bisa besar.

- Dataset lain, kondisi lingkungan berbeda, *hardware* berbeda, atau populasi lokasi lain.
- **Jika tetap unggul di domain baru → itu *novelty* kuat.**

3. Perbesar ukuran sampel (n)

Efekmu kecil? Bisa jadi **kurang daya uji**.

- Lakukan *power analysis* sederhana (boleh pakai AI/alat statistik).
- Targetkan **CI lebih sempit** dan **p-value** yang masuk akal (bukan dipaksa).

4. Lakukan *error analysis*

Jangan hanya lihat skor akhir. **Bedah kesalahannya.**

- *Bias vs variance*, kelas/segmen mana paling sering salah, kondisi uji mana rapuh.
- Temukan “kasus pembeda” → jadikan bahan eksperimen tambahan.

Mini-playbook iterasi (langkah cepat)

- **Pilih 1 target metrik** (mis. akurasi/F1/RMSE/efisiensi).
- **Skor dampak** tiap ide eksperimen:
 - *Impact* (0–3), *waktu/biaya* (0–3, terbalik), *risiko* (0–3, terbalik).
- **Ambil 3 ide dengan total skor tertinggi.**
- **Jalankan ringkas, catat angka, CI, dan kondisi uji.**
- **Update tabel perbandingan.** Jika $\Delta\%$ naik jelas → lanjutkan; jika tidak → ganti ide.

Contoh ide iterasi lintas bidang

- **Teknik/ML:** ganti *optimizer*, tambah *data augmentation*, *ablation* modul tertentu.
- **Material:** variasi rasio prekursor, suhu/tekanan, waktu sintesis; uji stabilitas.
- **Jaringan/sensor:** ubah interval sampling, posisi sensor, *firmware* versi baru.
- **Sosial/edukasi:** tambah durasi intervensi, *rubric* penilaian baru, segmen peserta.

Struktur hasil iterasi: angka → perbandingan → makna

- **Angka:** “RMSE turun dari **0,214 ±0,01** ke **0,192 ±0,01** (n=5 run).”
- **Perbandingan:** “**−10,3%** vs baseline internal; **−8,9%** vs Paper-B (protokol sama).”
- **Makna awal:** “Perbaikan konsisten pada *subset* berisik, menandakan ketahanan model.”

Template tabel ringkas (setelah iterasi)

Eksperimen	Variabel diubah	Domain/Populasi	Metrik (protokol)	Hasil (±CI)	Δ vs Baseline	Catatan
E1	Suhu 120→140° C	Sampel A	Yield, n=3 batch	72±2%	+9,1%	Stabil di batch kecil
E2	Dataset B (baru)	Domain B	F1, 5-fold, n=1200	0,842±.0 2	+6,5%	Robust domain shift
E3	Ablation modul “X”	A & B	RMSE, 3-run avg	0,192±.0 1	−10,3%	Modul X penting

Kotak Prompt #1 — rancang 3–5 eksperimen paling berdampak

“Tujuan: menguatkan novelty pada [metrik utama] di [domain].

Ini **kondisi awal**: [ringkas setup, angka, protokol].

Beri saya **5 ide eksperimen** dengan tabel (**Hipotesis** → **Variabel diubah** → **Desain uji** → **Sumber daya** → **Estimasi efek & risiko** → **Prioritas skor 0–3**).

Batasi pada yang bisa selesai < [durasi], biaya rendah, dan menjaga **apples-to-apples** dengan baseline.”

Bonus prompt (power & CI):

“Hitung kebutuhan **ukuran sampel (n)** untuk mendeteksi perbaikan relatif **≥[x]%** pada metrik [metrik], asumsi $SD=[...]$, $\alpha=0,05$, $power=0,8$. Sarankan pembagian fold/run.”

Kotak Prompt #2 — bandingkan temuan kita vs 3 paper teratas

“Bandingkan hasil kami [angka + protokol singkat] dengan 3 paper teratas berikut: [DOI-1], [DOI-2], [DOI-3].

Buat tabel apples-to-apples: Dataset/Domain, Metrik & Protokol, Nilai $\pm$ CI, $\Delta\%$ vs kami, Kondisi spesifik.

Lalu tulis 3 kalimat simpulan yang: (1) menyebut kapan kami unggul, (2) batasan yang tersisa, (3) langkah iterasi lanjutan yang paling logis.”

Tambahan validasi (anti-halu):

“Verifikasi DOI, halaman penerbit, dan cek apakah angka yang dikutip benar-benar ada (nomor tabel/gambar/halaman). Tandai preprint jika belum peer-reviewed.”

Tips cepat biar iterasimu tepat sasaran

- Fokus pada variabel berdampak tinggi. Jangan *random search*.
- Jaga protokol tetap konsisten saat membandingkan (itu *apples-to-apples*).
- Laporkan ketidakpastian: CI/SD/SE dan *n*.
- Simpan *audit trail*: *script*, versi alat, *seed*, tanggal run.
- Stop rule: jika 2–3 iterasi tidak menaikkan $\Delta\%$ secara bermakna → ganti tuas (pindah domain/variabel).

Ringkasan kilat

- Novelty lemah ≠ gagal. Artinya: butuh iterasi terarah.
- Empat tuas: variabel kritis, domain baru, ukuran sampel, error analysis.
- Desain 3–5 eksperimen dengan skor prioritas → jalankan yang paling menjanjikan.
- Bandingkan selalu dengan 3 paper teratas secara *apples-to-apples*.
- Tulis hasil dengan angka dan CI. Simpan semua sebagai jejak audit.
Hasilnya: *novelty* yang terukur, kredibel, dan siap tempur di jurnal.

Aksi 15 menit: pakai Kotak Prompt #1, dapatkan 5 ide eksperimen, pilih 3 teratas, lalu set *run* pertama. Besok, pakai Kotak Prompt #2 untuk memastikan kita benar-benar unggul.

Bagian IV — Mengkuantifikasi Klaim (Dari “A lebih baik” → “A 56% lebih baik”)

Metrik & Rumus Perbandingan

1. Dari “Kayaknya Baru” ke “Beneran Baru”

Sering kan, waktu riset atau skripsi, muncul rasa bangga karena merasa hasil kita **unik dan baru**? Tapi... tunggu dulu. Di dunia akademik, “*kayaknya baru*” **tidak cukup**. Kita harus bisa **membuktikan dengan data**.

Novelty bukan soal gaya menulis atau ide abstrak.

Novelty = *sesuatu yang belum pernah dicapai orang lain, dibuktikan dengan angka dan sumber valid*.

Bayangkan kamu bikin robot pengantar kopi. Kalau kamu bilang “robot saya lebih cepat”, reviewer pasti tanya:

“Seberapa cepat? dibanding siapa? pakai metrik apa?”

Nah, disinilah *novelty* bertemu dengan **angka**. Dan angka ini harus bisa diukur lewat **metrik kuantitatif**.

2. Saat Novelty Masih Lemah, Jangan Panik

Kadang hasil awal penelitian “flat”—tidak ada perbedaan signifikan. Itu bukan kegagalan. Itu **pintu awal untuk iterasi**.

Triknnya: **Ubah arah eksperimen secara cerdas**. Ada empat *tuas utama* untuk memperkuat *novelty*:

- **Ubah variabel kritis.**
Misal ganti *learning rate*, suhu, bahan, atau alat ukur. Perubahan kecil bisa membuat efek besar.
- **Coba domain baru.**
Uji di dataset atau konteks berbeda. Kalau hasil tetap bagus, berarti metode kamu **lebih robust**.
- **Perbesar ukuran sampel.**
Dengan data lebih banyak, hasil mu makin meyakinkan (dan bisa signifikan secara statistik).
- **Analisis kesalahan (*error analysis*).**
Lihat dimana sistem gagal. Dari situ, bisa muncul ide eksperimen baru yang lebih tajam.

3. Saatnya Berpikir Seperti Ilmuwan

Mahasiswa yang keren bukan cuma bisa coding atau bikin grafik.
Tapi bisa menjawab “**kenapa**” hasilnya seperti itu.

Coba renungkan:

- Apakah peningkatan hasil karena metode baru atau hanya karena dataset-nya berbeda?
- Apakah hasilmu masih berlaku jika diuji di konteks lain?
- Apakah perbedaan hasilnya besar secara **statistik** (*significant*) atau cuma kebetulan?

Jika kamu bisa menjawab itu semua dengan **angka**, berarti kamu sudah selangkah menuju paper yang kuat.

4. Mengubah Kalimat Kualitatif → Kuantitatif

Ini langkah penting banget. AI sering “membantu menulis,” tapi kamu tetap harus **menyediakan metrik nyata**.

Coba bandingkan:

✗ “Model A lebih baik dari Model B.”

✓ “Model A meningkatkan akurasi sebesar **+56%** dibanding Model B pada dataset X ($n=1200$, 5-fold CV).”

Simple, kan? Tapi dampaknya besar. Reviewer langsung tahu:

- Apa yang dibandingkan
- Berapa persen bedanya
- Kondisi uji-nya apa

5. Rumus Cepat untuk Mengkuantifikasi Klaim

Ada tiga rumus utama yang wajib kamu tahu dan hafal di dunia riset:

Relative Improvement

Mengukur peningkatan relatif antara A dan B.

$$\text{Relative improvement} = ((A - B) / B) \times 100\%$$

Contoh:

Model A = 0.89, Model B = 0.57

→ Relative improvement = $((0.89 - 0.57) / 0.57) \times 100\% = \mathbf{56\% \text{ lebih baik!}}$

Error Reduction

Mengukur penurunan kesalahan (lebih intuitif kalau metriknya “error”).

$$\text{Error reduction} = ((E_B - E_A) / E_B) \times 100\%$$

Misal error Model B = 0.20, Model A = 0.10

→ Error reduction = $((0.20 - 0.10) / 0.20) \times 100\% = 50\%$ **error berkurang.**

Effect Size (Cohen’s d)

Mengukur seberapa besar efek perubahanmu terhadap hasil. Nilai umum:

- $d = 0.2 \rightarrow$ kecil
- $d = 0.5 \rightarrow$ sedang
- $d = 0.8 \rightarrow$ besar

Selain itu, laporkan juga:

- **CI (Confidence Interval)** → rentang keyakinan hasil
- **p-value** → signifikansi statistik
- **F1/AUC/RMSE/MAE** → metrik sesuai bidangmu (AI, statistika, dll.)

6. Gunakan AI untuk Mengubah Narasi ke Angka

Kalau kamu menulis kalimat “A lebih cepat dari B”, minta AI bantu hitung ulang dengan *prompt*:

*“Ubah kalimat ini menjadi versi kuantitatif.
Sertakan metrik, $\Delta\%$, CI, dan n jika ada.”*

Contoh keluaran yang baik:

“Model A mencapai waktu proses 4,3 detik, **37% lebih cepat** dari Model B (6,8 detik, $n=10$ run, CI 95%).”

7. Tabel Cepat Perbandingan Kuantitatif

Studi	Dataset	Metrik	Nilai	Δ vs Baseline	Catatan
Paper-A (2023)	CIFAR-10	Akurasi	85.2 %	—	Baseline CNN
Paper-B (2024)	CIFAR-10	Akurasi	88.0 %	+3.3%	Transformer kecil
Kita (2025)	CIFAR-10	Akuras i	91.2 %	+7.3%	Performa tertinggi di dataset sama

Dari tabel ini kamu bisa langsung tulis klaim yang kuat:

“Metode kami meningkatkan akurasi sebesar **7,3%** dibanding Paper-B (2024) dengan **protokol uji identik**.”

8. Tantangan untuk Mahasiswa

Coba latihan ini:

1. Ambil hasil eksperimen sekarang (tabel atau grafik).
2. Hitung *relative improvement* dan *error reduction*.
3. Ubah satu paragraf hasilmu menjadi versi kuantitatif.
4. Lihat perbedaannya — tulisannya langsung terasa ilmiah, tajam, dan kredibel.

9. Kesimpulan Singkat

- **Novelty** tanpa angka = opini.
- **Novelty + metrik terukur = bukti ilmiah.**
- Gunakan rumus sederhana seperti *relative improvement* dan *error reduction* untuk memperkuat klaim.
- Latih AI untuk bantu menulis versi kuantitatif, tapi tetap **kamu yang validasi angka dan logikanya.**

Goal akhir:

Tulisanmu bukan cuma “kelihatan ilmiah,” tapi **punya angka yang bisa diuji ulang**. Dengan begitu, risetmu bukan sekadar tugas kampus — tapi **kontribusi nyata untuk ilmu pengetahuan**.

Membuat Pernyataan Terukur yang Kuat

Kenapa harus terukur?

Kalimat hebat itu **jelas** dan **terbukti**.

Bukan “A lebih baik dari B”, tapi “**A 56% lebih baik dari B**”.

Bukan “akurasi kita tinggi”, tapi “**akurasi naik 29% (CI 95%, n=1200)**”.

Intinya: klaim kuat = **angka + konteks uji + pembandingan**.

Komponen wajib dalam klaim

Saat menulis satu kalimat klaim, usahakan ada:

- **Metrik** yang spesifik: akurasi, F1, *RMSE*, *MAE*, waktu proses.
- **Pembandingan** yang jelas: *baseline* B, metode lama, atau paper tertentu.
- **Protokol uji**: *cross-validation 5-fold*, *holdout 20%*, *3-run average*.
- **Kondisi & data**: dataset, domain, alat, versi model.
- **Ukuran sampel (n)** dan **interval kepercayaan (CI)** jika tersedia.
- **Arah & besaran Δ** : “**+56%** dibanding B”.

Bold yang penting: metrik, $\Delta\%$, CI, n, dan konteks uji.

Istilah asing yang sering muncul: *baseline*, *protocol*, *cross-validation*, *effect size*, *confidence interval*, *statistically significant*.

Template kalimat siap pakai

Copy–edit saja sesuai angkamu.

- **Format umum**
“A [**metrik**]=X pada [dataset/domain], **Δ =Y%** vs *baseline* B, diuji dengan [*protocol*], n=..., **CI 95% ...**.”
- **Contoh 1 — “A 56% lebih baik dari B”**
“Model A meningkatkan F1 menjadi 0,89, **+56% dibanding Model B (0,57)** pada dataset X, *5-fold cross-validation*, n=1200, **CI 95% [0,87–0,91]**.”
- **Contoh 2 — “A 29% lebih akurat dari B”**
“Metode A mencapai akurasi 91,2%, **+29% dibanding B (70,7%)** di domain Y, *protocol holdout 20%*, n=1500, **CI 95% [89,9–92,4]**, perbedaan *statistically significant* ($p<0,05$).”
- **Contoh 3 — efisiensi waktu**
“A memotong waktu proses menjadi 4,3 detik, **–37% waktu vs B (6,8 detik)** pada mesin Z, *3-run average*, **CI 95% [4,1–4,5]**.”

Aturan emas kalimat klaim

- **Sebut metrik dan satuan.** Jangan “cepat”, tulis “detik”.
- **Sebut pembanding.** Tanpa B, $\Delta\%$ tidak bermakna.
- **Sebut *protocol*.** Biar *apple-to-apple*.
- **Laporkan n dan CI.** Ini bukti ketelitian, bukan hiasan.
- **Jangan klaim berlebihan.** Kalau beda kecil dan CI tumpang tindih, tulis “kecenderungan naik”, bukan “terbaik”.

Kesalahan umum (hindari!)

- “A lebih bagus” **tanpa metrik**.
- $\Delta\%$ **tanpa menyebut B** yang dibandingkan.
- Hasil di **dataset berbeda**, tapi ditulis seolah setara.
- Tidak mencantumkan **n/CI**, padahal datanya ada.
- Mencampur **akurasi** dan **F1** di satu klaim tanpa klarifikasi.

Mini-checklist 15 detik sebelum submit

- Sudah ada **metrik** dan **satuan**?
- Ada **baseline/paper pembanding** yang jelas?
- Tertulis **$\Delta\%$** atau selisih angka absolut?
- **Protocol uji** disebut (mis. *5-fold*, 3-run avg)?
- Ada **n** dan **CI** (kalau tersedia)?
- Konteks dataset/domain **jelas**?

Jika semua “ya”, klaimmu **kokoh**.

“A 56% lebih baik dari B” yang benar itu seperti apa?

Contoh kalimat final yang rapi:

“Pada dataset SensorX, *classifier A* mencapai $F1=0,89$, **+56%** dibanding *baseline B* (0,57). Evaluasi menggunakan 5-fold cross-validation, $n=1200$ sampel, CI 95% [0,87–0,91].”

Mengapa ini kuat?

- Ada **metrik (F1)**, **$\Delta\%$** , **baseline**, **protocol**, **n**, **CI**, dan **konteks dataset**.
- Reviewer bisa **menelusuri** dan **mengulang**.

“A 29% lebih akurat” yang tidak bikin debat

Contoh kalimat final yang rapi:

“Di domain Klinik-Y, model A mencapai akurasi 91,2%, **+29%** atas model B (70,7%) dengan *holdout* 20%, $n=1500$, CI 95% [89,9–92,4], perbedaan statistically significant ($p<0,05$).”

Kenapa mantap?

- Jelaskan **kapan** dan **bagaimana** diuji.
- Tunjukkan **ketidakpastian** (CI) dan **signifikansi** (p-value).
- Tidak ada ambiguitas.

Kotak Latihan (5 menit saja)

Ubah kalimat kualitatif di bawah ini menjadi versi kuantitatif yang kuat:

- “Model saya lebih bagus di dataset A.”
- “Metode baru lebih cepat dari script lama.”
- “A lebih stabil di domain berbeda.”

Target jawaban:

- Tambahkan metrik + angka A dan B.
- Tambahkan $\Delta\%$ dan *protocol*.
- Tambahkan n + CI (kalau ada).
- Tulis dalam **1 kalimat padat**.

Pro-tip pakai AI (tetap *data-first*)

Gunakan AI sebagai **editor angka**, bukan pengarang. Berikan tabel/grafik mentah, lalu minta AI **merapikan kalimat kuantitatif**. Contoh *prompt*:

“Ubah kalimat ini jadi klaim kuantitatif yang lengkap. Sertakan metrik, $\Delta\%$ terhadap baseline, *protocol* uji, n, dan CI 95% jika ada. Pastikan satu kalimat, padat, dan tidak berlebihan.”

Penutup singkat

- **Pernyataan terukur** membuat *novelty* terlihat nyata.
- **Angka + konteks + pembandingan** = klaim yang **tajam** dan **masuk akal**.
- Biasakan satu kalimat klaim yang **siap diaudit**.
- Ingat: **data dulu, narasi menyusul** — *data-first, anti-halusinasi*.

Visual Kuantitatif yang Meyakinkan

Kenapa visual penting?

Bayangkan kamu sedang presentasi hasil penelitian di depan dosen atau reviewer. Kamu berkata, “*Model saya 56% lebih baik dari baseline.*” Tanpa visual, pernyataan itu cuma **angka kering**. Tapi kalau kamu tunjukkan grafik dengan perbandingan, **pesanmu langsung nempel di otak**.

Visual yang baik bisa menjelaskan 1 halaman teks hanya dalam 1 detik.

Makanya, bagian ini fokus ke **bagaimana bikin grafik dan tabel yang “berbicara”**.

Ciri-ciri visual yang meyakinkan

Visual yang kuat bukan yang penuh warna dan efek, tapi yang **mudah dibaca dan bisa diverifikasi**. Ada beberapa elemen penting:

- **CI (Confidence Interval)** → menunjukkan tingkat kepercayaan hasil. Misal, “akurasi $91\% \pm 1,5\%$ ” artinya ada variasi kecil tapi hasilnya konsisten.
- **SE (Standard Error)** → memperlihatkan seberapa stabil model saat diulang.
- **Ablation Bars** → tunjukkan hasil jika satu komponen dihapus. Misalnya, “kalau modul X dihilangkan, performa turun 8%”. Ini bukti bahwa setiap bagian model punya peran nyata.
- **Δ Table (Delta Table)** → tabel sederhana yang memperlihatkan *selisih performa* antar metode. Δ = perubahan/peningkatan yang bisa diukur.

Bentuk visual yang paling efektif untuk laporan atau jurnal

a. Grafik batang (bar chart)

Cocok untuk menampilkan *delta performance* antar metode. Tambahkan *error bar* (CI/SE) biar pembaca tahu seberapa yakin datanya.

b. Grafik garis (line chart)

Bagus untuk menunjukkan tren, misal akurasi terhadap jumlah *epoch* atau ukuran data. Gunakan warna berbeda untuk tiap model, tapi tetap sederhana.

c. Tabel perbandingan kuantitatif

Tabel tetap penting karena reviewer suka angka mentah. Gunakan format seperti ini

Studi	Dataset	Metrik	Nilai	Δ vs Baseline
Paper A (2023)	CIFAR-10	Akurasi	88.0%	—
Paper B (2024)	CIFAR-10	Akurasi	90.5%	+2.5%
Kita (2025)	CIFAR-10	Akurasi	91.2%	+3.2%

Tips: Gunakan **bold** untuk hasil terbaik, dan jangan takut menampilkan $\Delta\%$ secara jujur. Reviewer lebih menghargai transparansi daripada klaim berlebihan.

Panduan praktis bikin visual

- Gunakan **warna konsisten** antar grafik (misal, model kamu = biru, baseline = abu-abu).
- Tambahkan **label angka langsung di atas batang**, agar pembaca tak perlu menebak.
- Hindari **3D chart** — terlihat keren tapi bikin data susah dibaca.
- Jika pakai AI untuk bantu visualisasi (misal dari Python/Excel), tetap periksa **angka CI & n**.
- Simpan semua data mentah di *appendix audit* untuk menjaga *traceability*.

Ablation study = bukti kamu paham sistemmu

Banyak mahasiswa cuma fokus pada hasil akhir, padahal reviewer ingin tahu **mengapa** hasil itu bisa terjadi. Nah, di sinilah *ablation study* berguna.

Contoh:

Kamu punya model dengan tiga modul (A, B, C). Coba jalankan tiga variasi eksperimen:

- Tanpa A
- Tanpa B
- Tanpa C

Lalu tampilkan grafik batang:
semakin banyak modul dihapus → semakin turun performa.
Itu **bukti ilmiah** bahwa tiap bagian penting.

Reviewer suka melihat *ablation bar chart* karena menunjukkan *depth of analysis*, bukan sekadar hasil akhir.

Tabel *Delta* (Δ) = bahasa universal antar peneliti

Kamu bisa pakai *delta table* sederhana untuk semua bidang: AI, biologi, pendidikan, ekonomi, dll. Strukturnya:

Metode	Metrik	Nilai ($\pm$ CI)	Δ vs Baseline	Catatan
Model B (baseline)	Akurasi	70.7% $\pm$ 1.2	—	Standar lama
Model A (baru)	Akurasi	91.2% $\pm$ 1.0	+29%	Peningkatan signifikan
Model A tanpa modul X	Akurasi	83.5% $\pm$ 1.4	-8%	Modul X krusial

Tabel seperti ini memudahkan pembaca langsung **melihat tren, perbedaan, dan kontribusi tiap eksperimen**.

Tips terakhir: “Grafik bukan dekorasi”

Jangan gunakan grafik hanya karena ingin terlihat “ilmiah”. Gunakan grafik untuk **menceritakan sesuatu**. Kalau sebuah gambar tidak menambah makna, hapus saja.

Ingat prinsipnya: “**Less decoration, more interpretation.**”

Kesimpulan

- **Visual = senjata utama** untuk membuat data berbicara.
- Gunakan **grafik dengan CI/SE** untuk menunjukkan keandalan hasil.
- Tambahkan **ablation bars** untuk membuktikan kontribusi tiap komponen.
- Gunakan **Δ table** untuk menegaskan perbedaan performa secara numerik.
- Buat visualmu **sederhana, jelas, dan bisa diaudit**.

Tantangan untuk kamu hari ini:

Ambil satu grafik dari hasil eksperimen.

Tambahkan CI/SE dan satu bar “tanpa modul utama”.

Lalu ubah jadi delta table.

Kamu akan langsung merasa risetmu lebih solid, meyakinkan, dan siap publikasi.

Bagian V — Menyusun Kesimpulan → Abstrak → Pendahuluan (Urutan Terbalik yang Efisien)

Kesimpulan Terarah Aksi (Actionable)

Bayangkan kesimpulanmu seperti *pitch* 30 detik. Cepat. Padat. **Ada angka**. Ada **dampak**. Bukan cerita panjang.

Prinsip utama:

- **3–5 butir** saja. Biar fokus.
- Setiap butir **ada metrik + angka + kondisi uji**.
- Tulis **apa artinya** bagi pengguna/peneliti/industri.
- Cantumkan *traceability*: **(lihat Tabel/Gambar X)**.

Anatomi 1 butir kesimpulan

[Apa yang naik/turun] + [Angka inti $\pm$ CI, $\Delta\%$ vs *baseline*] + [Kondisi uji] + [Dampak nyata] (lihat Tabel/Gambar X)

Contoh format kalimat:

- **Akurasi naik 6,8% (CI 95% $\pm$ 1,2)** pada **Dataset C** dibanding *baseline* B, **menurunkan kesalahan labeling harian $\sim$ 1.400 sampel** di lini produksi. (*lihat Tabel 2*)

Contoh 4 butir (siap pakai)

1. **Model A lebih akurat 29%** dari *baseline* B (**91,2% $\pm$ 1,0 vs 70,7% $\pm$ 1,2**), uji **5-fold**, **n=1.200**. **Dampak:** kualitas deteksi naik; **post-editing** berkurang. (*lihat Gambar 3*)
2. **Waktu inferensi turun 41%** pada GPU kelas menengah. **Dampak:** **biaya komputasi** per 1.000 prediksi turun signifikan. (*lihat Tabel 4*)
3. **Robust di domain baru:** performa hanya turun **2,1%** saat *domain shift* B→C. **Dampak:** rilis lintas-klien tanpa retuning panjang. (*lihat Tabel 5*)
4. **Ablation:** menghapus **modul X** menurunkan F1 **–8%**. **Dampak:** modul X **wajib** dipertahankan pada rilis produksi. (*lihat Gambar 5*)

Catatan gaya: pakai istilah *baseline*, *domain shift*, *ablation*, *CI* dengan miring. Sorot angka **tebal** agar “nendang”.

Do & Don't singkat

Do

- **Angka dulu, tafsir sesudah.**
- Tulis **kondisi uji**: dataset, protokol, n , *seed* bila ada.
- Kaitkan ke **dampak**: biaya, waktu, akurasi lapangan, keamanan, UX.

Don't

- Hindari “**lebih baik**” tanpa angka.
- Jangan generalisasi di luar **kondisi uji**.
- Jangan lupa *traceability* ke **tabel/gambar**.

Batasan (Limitations) & Arah Kerja Lanjut

Kesimpulan yang matang **jujur pada batasan**. Ini bukan kelemahan. Ini **peta jalan** risetmu.

Cara menulis batasan yang sehat

- Sebut **ruang uji**: dataset, ukuran sampel, perangkat, *hyperparameter*.
- Jelaskan **risiko bias**: *class imbalance*, *selection bias*, *overfitting*.
- Nyatakan **ketidakpastian**: **CI lebar**, *variance* tinggi, hasil sensitif *seed*.

Contoh kalimat:

- **Batasan 1** — *External validity*: uji hanya di **Dataset C**; domain **low-light** belum diuji.
- **Batasan 2** — *Compute constraint*: uji latensi pada **1 jenis GPU**; CPU edge belum dievaluasi.
- **Batasan 3** — *Data imbalance*: kelas minor <5%, **F1 minor** masih fluktuatif ($\pm 2,8$).

Arah kerja lanjut yang spesifik (bukan generik)

Gunakan pola **aksi** → **metrik** → **target** → **kondisi**.

Contoh:

- **Perluas domain** ke data **low-light** dan **noisy audio**; target $\Delta F1 \leq -1\%$ terhadap domain asli, $n \geq 2.000$.
- **Skalakan ukuran sampel 2×** untuk mempersempit **CI** akurasi menjadi $\pm 0,7$ pada *5-fold*.
- **Optimasi CPU edge**: capai **<40 ms/inferensi** pada **Intel i5 gen 11**, batch=1.
- **Regularisasi minor class** (focal loss, *data augmentation* sintetis) agar **F1 minor +5%** tanpa menurunkan akurasi mayor >1%.

- **Uji stress:** *seed sweep* (10 seed) + **sensitivity analysis learning rate** untuk menilai stabilitas.

Mini-Template siap tempel

Template “Kesimpulan Terarah Aksi” (isi angka kalian)

1. **[Metrik utama]** naik/ turun **[X%]** ($\pm$ CI) pada **[Dataset/Domain]** vs *baseline* **[Nama]**; dampak **[biaya/waktu/kualitas]**. (*lihat [Tabel/Gambar]*)
2. **[Efisiensi]:** latensi **[ms]**, throughput **[req/s]**, biaya/1k prediksi turun **[Y%]**; tetap stabil di **[perangkat]**. (*lihat [Tabel]*)
3. **[Robustness]:** *domain shift* **[A→B]** menurunkan **[metrik]** hanya **[Z%]**; **impikasi rilis:** **[apa]**. (*lihat [Gambar]*)
4. **[Ablation]:** modul **[X]** menyumbang **[Δ metrik]**; **rekomendasi:** pertahankan/ ganti komponen **[alasan singkat]**. (*lihat [Gambar]*)

Template “Batasan & Arah Lanjut”

- **Batasan** — **[aspek]:** diuji pada **[dataset/perangkat/n kecil]**, ada **[bias/CI lebar/sensitivitas]**.
- **Arah lanjut 1** — **[aksi teknis]** untuk **[metrik]**, target **[angka]**, kondisi **[protokol]**.
- **Arah lanjut 2** — **[aksi eksperimen]** pada **[domain/ukuran sampel]** agar **[stabilitas/CI]** tercapai.
- **Arah lanjut 3** — **[optimasi resource]** demi **[latensi/biaya/energi]** di **[hardware]**.

Checklist 45-detik sebelum submit

- Ada **3–5 butir** kesimpulan.
- Setiap butir punya **angka + CI/ Δ + kondisi uji**.
- Ada **dampak nyata** (biaya/waktu/kualitas/risiko).
- **Batasan** ditulis spesifik, bukan umum.
- **Arah lanjut** berbentuk **aksi terukur** (ada target & protokol).
- Semua butir punya **rujukan ke tabel/gambar** (*traceability*).

Tantangan singkat

Ambil hasil terbaru. Tulis **4 butir kesimpulan** pakai template di atas. Tambahkan **3 batasan + 3 aksi lanjut** yang bisa kamu eksekusi minggu ini. Rasakan bedanya: tulisanmu jadi **tajam, akuntabel, dan siap dilawan reviewer**.

Abstrak yang Tajam & *Critical* (Highlight Novelty)

Pernah menulis abstrak tapi dosen bilang:

“Ini belum tajam, masih deskriptif.”

Tenang. Hampir semua mahasiswa pernah di situ. Masalahnya sederhana: **abstrakmu belum punya alur logika ilmiah yang jelas**. Formula ajaibnya sebenarnya cuma lima langkah:

gap → tujuan → metode → angka kunci → dampak.

Mari kita bedah satu per satu.

1. Gap — Lubang Pengetahuan yang Kamu Isi

Mulailah dengan **masalah nyata** yang belum terselesaikan. Bukan sekadar “belum banyak yang meneliti,” tapi **apa yang belum efektif, belum efisien, atau belum akurat** dari studi sebelumnya.

Contoh:

“Model deteksi emosi pada teks Bahasa Indonesia masih rendah akurasi di domain media sosial karena banyak *slang* dan *emoji*.”

Tips:

- 1 kalimat cukup.
- Gunakan kata pemicu seperti: *namun, masih rendah, belum optimal, kurang efisien*.
- Jangan pakai kalimat panjang yang menjelaskan sejarah topik.

2. Tujuan — Apa yang Kamu Lakukan untuk Menambal Gap Itu

Setelah menyebut masalah, langsung tunjukkan solusinya.

Gunakan gaya kalimat yang lugas:

“Penelitian ini mengusulkan model BERT-ID yang dioptimasi dengan *domain-specific pretraining* untuk meningkatkan akurasi klasifikasi emosi.”

Kuncinya:

Kalimat ini menjawab “**apa yang kamu tambahkan baru**”. Inilah bagian awal *novelty*-mu!

3. Metode — Bagaimana Kamu Mencobanya

Bagian ini bukan laporan lengkap, tapi cukup **satu kalimat ringkas** tentang pendekatan utama.

Contoh:

“Metode diuji menggunakan dataset 12.000 komentar media sosial dengan *cross-validation* 5-fold dan perbandingan terhadap tiga model *baseline*.”

Tips singkat:

- Sebutkan *dataset*, *metode uji*, dan *baseline*.
- Hindari daftar alat (“kami menggunakan Python, Excel, TensorFlow...”). Reviewer tidak butuh itu di abstrak.

4. Angka Kunci — Bukti Nyata *Novelty*-mu

Di sinilah kamu tunjukkan *impact*. Gunakan **angka dan $\Delta\%$ (delta)** dibanding *baseline*.

Contoh:

“Model ini mencapai akurasi 91,2% ($\Delta=+29\%$ dibanding baseline BiLSTM) dengan CI 95% $\pm 1,0$.”

Kalimat ini membuat reviewer langsung berpikir:

“Wah, ini signifikan. Ada angka, ada pembandingan, ada kondisi uji.”

Tanpa angka, abstrak lemah.

Dengan angka, abstrak bicara.

5. Dampak — So What?

Terakhir, jelaskan **kenapa hasilmu penting**. Bukan “untuk penelitian selanjutnya,” tapi efeknya di dunia nyata atau akademik.

Contoh:

“Temuan ini menunjukkan potensi integrasi *domain-adaptive language model* dalam analisis sentimen Bahasa Indonesia, terutama untuk platform publik dan aplikasi kebijakan digital.”

Satu kalimat, tapi punya **nilai aplikasi dan kontribusi teoretis**.

Dua Gaya Abstrak yang Bisa Kamu Pilih

A. Fokus Aplikasi (buat riset terapan)

- Mulai dari **masalah praktis**.
- Tekankan hasil yang **bisa langsung digunakan**.
- Tutup dengan dampak **bagi masyarakat, industri, atau lembaga**.

Contoh gaya kalimat:

“Pendekatan ini menurunkan waktu analisis data transaksi dari 15 menit menjadi 3 menit, mendukung otomatisasi audit keuangan UMKM.”

B. Fokus Kontribusi Teoretis (buat riset metodologis)

- Tekankan **konsep baru atau framework**.
- Soroti **perbandingan matematis, metrik, atau pendekatan** yang lebih baik.
- Tutup dengan implikasi ke arah riset lanjutan.

Contoh:

“Model regularisasi baru ini memperkuat stabilitas *loss landscape* dan memperbaiki *generalization* pada dataset terbatas, membuka arah baru untuk riset *low-resource learning*.”

Kotak Prompt (Gunakan AI untuk Bantu Menyusun Abstrak)

Kalau kamu sudah punya tabel hasil & kesimpulan, kamu bisa minta bantuan AI dengan *prompt* seperti ini:

Prompt siap pakai:

“Tulis abstrak 200 kata dari tabel hasil berikut. Gunakan struktur *gap* → *tujuan* → *metode* → *angka kunci* → *dampak*. Fokuskan pada *novelty* dan buat dua versi: satu bergaya aplikasi, satu bergaya teoretis.”

AI akan bantu merapikan, tapi tetap kamu yang harus memastikan **semua angka valid** dan **tidak ada halusinasi**.

Checklist Sebelum Kamu Kirim Abstrak

- Panjang 150–250 kata.
- Ada *gap* (masalah jelas).
- Tujuan dijelaskan dalam 1 kalimat kuat.
- Metode disebut ringkas tapi jelas.
- Ada angka kunci (metrik + $\Delta\%$ vs *baseline*).
- Ada dampak (praktis atau teoritis).
- Tidak ada klaim tanpa data.

Tantangan untuk Kamu

Coba ubah abstrak lamamu. Hitung kata, pastikan ada 5 unsur tadi. Tambahkan **1 angka kunci**, **1 kalimat dampak nyata**, dan **1 kalimat pembeda dari riset lama**. Lihat perbedaannya — abstrak jadi **tajam, kritis, dan langsung terasa “ilmiah”**.

Ingat: Abstrak bukan sinopsis. Itu adalah “iklan akademik” dari karya ilmumu. Satu paragraf yang menentukan apakah orang mau membaca seluruh jurnalmu atau tidak. Jadi buatlah abstrak **tajam seperti pisau, berbasis data, dan anti-halusinasi**.

Pendahuluan Tanpa Halusinasi

Pendahuluan adalah “pintu depan” paper-mu. Kalau pintunya kokoh, pembaca lanjut. Kalau rapuh, mereka balik. Kokohnya pendahuluan datang dari **referensi yang valid** dan **alur logika yang rapi**. Target kita: **tanpa halusinasi**, **berbasis data**, dan **adil terhadap karya orang lain**.

Teknik Cari Referensi (Google Scholar, Semantic Scholar, *Elicit*)

Tujuan: cepat dapat sumber **asli**, **terverifikasi**, dan **relevan**.

A. Google Scholar

- Ketikkan kueri yang spesifik: “**subword**” **Indonesian sentiment** “**F1 score**” **2019..2025**
- **Filter**: tahun 5 tahun terakhir.
- **Klik penerbit** (ACM/IEEE/Elsevier/Springer/MDPI/ArXiv).
- **Cek DOI**. Kalau ada halaman penerbit + PDF resmi, **mantap**.
- **Hindari**: blog, *preprint* tanpa konteks, kutipan yang hanya menyalin judul.

B. Semantic Scholar

- Lihat “**Highly Influential Citations**” dan “**Cited By**”.
- Manfaat: mudah telusuri *forward/backward citations* (siapa mengutip & siapa yang dikutip).
- **Baca abstrak + figures** untuk menilai cepat kontribusi.

C. *Elicit*

- Tulis pertanyaan: “What improves Indonesian sentiment classification under domain shift?”
- Dapatkan ringkasan singkat + **metode + metrik + dataset**.
- **Validasi ulang** di Scholar/penerbit. *Elicit* itu asisten, bukan wasit terakhir.

Checklist Anti-Halusinasi (pakai setiap kali):

- **Ada DOI?**
- **Ada halaman penerbit?**
- **Tahun sesuai fokus (≤ 5 tahun)?**
- **Metrik/angka jelas (F1/AUC/RMSE/CI)?**
- **Bisa diunduh dan dibaca penuh (bukan hanya judul)?**

Ingat: AI boleh bantu mencari, **kamu** yang wajib memverifikasi.
Trust, but verify.

Menjahit *Background* → *Research Gap* → *Objectives* Berbasis Data

Pendahuluan yang baik itu **mengalir**: dari gambaran umum, ke kekurangan, lalu ke tujuanmu.

Pola 3-paragraf yang aman:

1. **Background (ringkas, faktual)**
 - Kenalkan konteks + pentingnya masalah.
 - Cantumkan **2–3 referensi inti** (bukan 10).
 - Hindari cerita sejarah panjang.
2. **Research Gap (tajam, spesifik)**
 - Jelaskan **apa yang belum berhasil**: akurasi rendah, *domain shift*, biaya komputasi tinggi, *sample size* kecil, dsb.
 - Sertakan **angka dari literatur**: “F1 0,78 pada domain X (n=12k).”
 - Tunjukkan **keterbatasan protokol**: “tanpa *cross-domain validation*,” “tidak ada *ablation*,” dll.
3. **Objectives (terukur, singkat)**
 - “Paper ini **mengusulkan** ...”
 - “Kami **mengevaluasi** pada dataset ... dengan **protokol** ...”
 - “Kami **membandingkan** terhadap 3 *baseline* ... dan melaporkan **CI 95%**.”

Template cepat (silakan *copy–edit*):

- *Background*:
“Klasifikasi emosi berbahasa Indonesia penting untuk moderasi dan layanan publik. Metode X dan Y telah dilaporkan efektif pada domain berita.”
- *Gap*:
“Namun, kinerja di media sosial masih **rendah (F1 0,78)** karena *slang* dan *emoji*, serta belum ada evaluasi **lintas domain** dan **ablation** yang jelas.”
- *Objectives*:
“Kami **mengusulkan** fine-tuning BERT-ID dengan *domain-adaptive pretraining*. Kami **mengevaluasi** pada 12k komentar (5-fold, *cross-domain*), **membandingkan** dengan BiLSTM/fastText/SVM, dan **melaporkan** F1 + **CI 95%**.”

Kunci: angka dari literatur di bagian *gap*, dan **rencana metrik/jumlah data** di bagian **objectives**.

Related Work yang Adil & Apples-to-Apples

“Adil” artinya kamu **tidak cherry-pick**.

“*Apples-to-apples*” artinya **bandingkan pada metrik & protokol yang sama**.

Prinsip singkat:

- **Kelompokkan** karya berdasarkan *metode* atau *domain*.
- **Sebutkan metrik + protokol** yang dipakai setiap studi.
- **Tulis keterbatasan** dengan netral: “Studi A kuat di domain berita, **belum diuji** di media sosial.”
- **Jangan klaim menang** jika dataset/metrik beda. Tulis “tidak langsung sebanding.”

Mini-rubrik “adil”:

- Referensi inti 3–7 studi, 5 tahun terakhir.
- Ada angka (bukan deskripsi doang).
- Ada *protocol note* (fold, *train/test split*, *cross-domain*).
- Ada *limits note* (kapan kuat, kapan lemah).
- Tidak ada klaim di luar data.

Kotak Prompt (bangun paragraf *related work* yang adil)

Prompt #1 — Ringkas *related work* anti-halusinasi

“Ringkas 5 paper teratas tentang [topik]. Untuk setiap paper, tulis: *metode*, *dataset*, *metrik*, *angka kunci*, *protokol uji*, *keterbatasan*. Jangan berhalusinasi; kutip DOI & penerbit. Jika tidak yakin, tulis ‘tidak tersedia di sumber asli’.”

Prompt #2 — Susun paragraf *apples-to-apples*

“Dari ringkasan tadi, tulis satu paragraf *related work* yang adil. Kelompokkan metode, bandingkan hanya pada metrik & protokol yang sama. Hindari klaim superioritas jika dataset beda. Tutup dengan celah yang masih ada.”

Prompt #3 — Verifikasi silang

“Buat daftar DOI + tautan penerbit untuk 5 studi yang digunakan. Tandai mana yang *preprint* dan mana yang peer-reviewed. Jika ada metrik yang tidak dilaporkan, beri catatan.”

Prompt #4 — Red flag detektor

“Periksa paragraf *related work* ini. Tandai kalimat yang berpotensi halusinasi, referensi tanpa DOI, atau klaim ‘mengungguli’ yang tidak *apples-to-apples*. Usulkan perbaikan.”

Do & Don’t (super singkat)

Do:

- **Verifikasi DOI dan baca PDF-nya.**
- **Tulis angka + metrik + protokol** saat membahas studi.

- Tunjukkan **celah** yang nyata, bukan opini.

Don't:

- Mengutip “katanya” tanpa sumber asli.
- Membandingkan metrik beda dataset lalu klaim menang.
- Menyebut alat/versi AI tanpa **disclosure** dan **validasi**.

Alur 30-Detik Sebelum *Submit* Pendahuluan

1. Ada **3–7** referensi utama dengan **DOI**.
2. *Background* ≤ 5 kalimat, **to the point**.
3. *Gap* punya **angka** dari literatur.
4. *Objectives* punya **metrik + protokol** yang akan kamu pakai.
5. *Related work* **adil** dan *apples-to-apples*.

Jika kelima poin ini **centang**, pendahuluan **anti-halusinasi** dan siap jalan.

Intinya: gunakan alat cerdas (Google Scholar, Semantic Scholar, *Elicit*) **untuk menemukan**, lalu gunakan **otak ilmiahmu untuk memverifikasi**. Pendahuluan jadi kuat, objektif, dan **siap meyakinkan reviewer**.

Bagian VI — Tinjauan Pustaka & Perbandingan dengan Studi Lain

Mencari Referensi yang Tidak Halusinasi

Tujuan kita sederhana: **dapat sumber asli, valid, dan bisa dipercaya.**

Kita lawan *halusinasi* dengan **verifikasi** yang disiplin.

Ingat: **judul keren ≠ bukti kuat.**

Tiga Kunci Verifikasi: DOI, Jurnal/Indeks, *Publisher Page*

1) Cek DOI (Digital Object Identifier)

- DOI itu **ID unik** sebuah artikel.
- Kalau ada DOI, kita bisa **lacak ke sumber resmi**.
- **Tidak ada DOI?** Hati-hati. Bisa jadi *preprint* atau sumber kurang jelas.

2) Cek Jurnal & Indeks

- Apakah jurnal **terindeks** (mis. Scopus, Web of Science)?
- Ada **proses peer-reviewed?**
- Lihat **scope** jurnal. Cocok tidak dengan topikmu?
- **Hindari** jurnal yang tidak jelas *editorial board*-nya.

3) Cek *Publisher Page*

- Buka halaman **penerbit resmi** (*publisher page*).
- Pastikan ada **PDF resmi, detail volume/issue, dan tanggal**.
- Bandingkan dengan versi *preprint* jika ada perbedaan.

Inti: DOI valid + jurnal kredibel + *publisher page* = referensi kuat.

Hindari “Judul Terlihat Ilmiah” Tanpa Rekam Jejak

Banyak tulisan **tampak ilmiah**, tapi **buktinya tipis**. Jangan terkecoh.

Tanda-tanda *red flag* (waspada):

- Tidak ada **DOI** atau DOI tidak bisa dilacak.
- **Jurnal baru** tanpa indeks jelas, tapi klaimnya bombastis.
- **Tidak ada metrik/angka** (mis. F1/AUC/CI).
- Klaim “**mengungguli semua**” tapi **dataset/metrik tidak jelas**.
- PDF hanya *slide, blog, atau whitepaper* marketing.
- *Publisher* tidak jelas, domain web aneh, atau **konten duplikasi**.

Tanda-tanda *green flag* (aman):

- Ada **DOI** yang aktif.

- Jurnal terindeks, jelas **proses peer-reviewed**.
- Ada **metode, dataset, metrik, dan angka** lengkap.
- Ada **data/code availability** atau lampiran.
- Sitasi masuk akal, **5 tahun terakhir** dominan.

Alur 60-Detik Anti-Halusinasi (Praktis & Cepat)

1. **Cari judul** di Google Scholar / Semantic Scholar.
2. **Klik DOI** → buka **halaman penerbit**.
3. Catat **jurnal, tahun, volume/issue**.
4. Lihat **metrik/angka** (mis. RMSE 0,19; F1 0,84 ±0,02).
5. Konfirmasi **indeks** (Scopus/WoS) atau minimal reputasi penerbit.
6. Simpan **screenshot/URL penerbit** di *appendix audit*.

Kalau 1 langkah gagal, jangan masukkan referensi itu ke *literature review*.

Contoh Mini-Audit (cepat ditempel di catatan lab)

- **DOI:** 10.xxxx/xxxxx (aktif  / mati )
- **Jurnal & Indeks:** Nama jurnal; Scopus/WoS (ya/tidak)
- **Publisher page:** ada (link)
- **Peer-review:** ya / tidak jelas
- **Metrik & angka kunci:** F1=0,842 ±0,02 (n=1200, 5-fold)
- **Catatan:** metode jelas / dataset publik / ablation ada

Tips Kuat Biar Nggak Tersesat

- **Selalu bandingkan 2–3 sumber** untuk topik yang sama.
- **Prioritaskan peer-reviewed**; *preprint* boleh, **tapi** verifikasi ekstra.
- Baca **bagian metode** dulu. Kalau metode kabur, **skip**.
- Cek **sitasi silang**: siapa mengutip, dan dikutip siapa.
- Simpan **PDF resmi** dan **kutip halaman** saat mengambil angka.

Latihan Cepat (5 Menit)

- Ambil 1 artikel yang kamu pakai.
- Isi mini-audit di atas.
- Jika **ada 2 red flag** → ganti referensi.
- Jika **semua hijau** → pakai, dan tandai sebagai **sumber utama**.

Kotak Mini-Prompt (opsional, untuk bantu verifikasi)

“Ringkas artikel ini hanya dari *publisher page*/PDF resmi. Tulis: **DOI, jurnal, tahun, metrik & angka, protokol uji, dan apakah *peer-reviewed***. Tandai bagian yang tidak jelas. Jangan berhalusinasi.”

“Bandingkan dua artikel berikut. Pastikan **metrik & protokol** benar-benar sama sebelum menyimpulkan ‘lebih baik’. Jika tidak *apples-to-apples*, beri catatan peringatan.”

Ringkasan Super Pendek

- **Verifikasi DOI** → pastikan **asli & aktif**.
- **Cek jurnal/indeks** → ada *peer-review* dan reputasi.
- **Buka *publisher page*** → cocokkan detail & unduh PDF resmi.
- **Hindari** referensi yang hanya “kelihatan ilmiah” tanpa bukti.
- **Simpan jejak verifikasi** untuk audit.

Dengan langkah ini, *literature review* kamu jadi **bersih, tajam, dan anti-halusinasi**. Reviewer pun senyum.

Membandingkan Hasil Kita dengan Literatur

Tujuan kita jelas: **buktikan posisi hasil kita** terhadap karya orang lain. Bukan sekadar “bagus”, tapi **terukur** dan **adil**. Kuncinya: **tabel komparasi** yang rapi dan *apples-to-apples*.

Prinsip Utama

- **Bandungkan hal yang sama: dataset, metrik, protokol uji** harus sejalan.
- **Tulis angka, bukan opini:** sertakan **nilai, $\pm$ CI/SE**, dan **n**.
- **Hitung selisih ke SOTA** (*state-of-the-art*): tampilkan **$\Delta\%$** dengan jelas.
- **Catat konteks teknis:** versi data, *preprocessing*, *hardware* bila relevan.

Ingat: tanpa kesamaan protokol, perbandingan tidak sah.

Tabel Komparasi: Kolom Wajib

Gunakan format ringkas ini:

- **Studi** → nama/kode paper atau *author–year*.
- **Dataset** → nama + *split* (mis. train/val/test, atau *5-fold*).
- **Metrik (protokol)** → contoh: **F1, 5-fold, n=1200**.
- **Hasil ($\pm$ CI)** → contoh: **0,842 $\pm$ 0,02**.
- **Δ vs SOTA** → persen naik/turun relatif ke baris SOTA.
- **Catatan** → alasan teknis: *augmentation*, *ablation*, *domain shift*, dll.

Tips: jadikan baris **SOTA tebal** agar kontras. Baris kita diberi ikon .

Cara Hitung Δ vs SOTA (singkat)

- $\Delta\% = ((\text{Skita} - \text{SSOTA}) / \text{SSOTA}) \times 100\%$ untuk metrik “makin besar makin baik” (mis. F1, AUC, accuracy).
- $\Delta\% = ((\text{SSOTA} - \text{Skita}) / \text{SSOTA}) \times 100\%$ untuk metrik “makin kecil makin baik” (mis. RMSE, MAE, latensi).

Tuliskan tanda $\pm$ jika ada CI/SE. Contoh: **+6,5% ($\pm$ 1,2%)**.

Checklist “Apples-to-Apples”

- **Dataset sama?** Nama, versi, *split* konsisten.
- **Metrik sama?** Definisi F1 (macro/micro), **AUC**, **RMSE** jelas.
- **Protokol sama?** *k-fold*, jumlah *run*, *seed* acak, *early-stopping*.
- **Preprocessing sama?** Normalisasi, *augmentation*, *tokenization*.
- **Sumber daya tercatat?** *Hardware*, waktu latih, *params* utama.
- **Statistik dilaporkan?** $\pm$ CI/SE, **n**, atau **3-run avg**.
- **Ablation ada?** Jelaskan bagian model yang bikin naik/turun.

Kalau ada beda protokol, beri **catatan peringatan** di kolom **Catatan**.

Mengapa Bisa Beda? (Penyebab Teknis)

- Distribusi data beda (*domain shift*).
- Preprocessing tidak identik.
- Hyperparameter dan *seed* berbeda.
- Implementasi atau versi *library* beda.
- Data leakage pada *split* atau *tuning*.
- Overfitting karena *early-stopping* tak konsisten.

Tuliskan penyebab ini secara jujur di kolom **Catatan**. Reviewer suka transparansi.

Contoh Ringkas Tabel (isi sesuai studi kalian)

Studi	Dataset	Metrik (protokol)	Hasil ($\pm$ CI)	Δ vs SOTA	Catatan
SOTA (Doe-2024)	CIFAR-10 (std test)	Accuracy, 5-run	96,8% $\pm 0,2$	—	Augmentation kuat
Kita (✓)	CIFAR-10 (std test)	Accuracy, 5-run	97,2% $\pm 0,3$	+0,41%	Regularization + mixup
Lee-2023	CIFAR-10 (std test)	Accuracy, 3-run	96,1% $\pm 0,4$	-0,72% vs SOTA	Ablation modul X

Catatan: pastikan “std test” berarti *split* publik yang sama.

Narasi Pendek Setelah Tabel (format 3 kalimat)

1. **Angka:** “Model kami meraih **97,2% $\pm 0,3$** pada CIFAR-10 (5-run).”
2. **Perbandingan:** “Ini **+0,41%** di atas **SOTA** (96,8% $\pm 0,2$).”
3. **Makna teknis:** “Kenaikan terutama dari **mixup** dan **regularization**; *augmentation* kami setara SOTA.”

Urutan sakti: angka → perbandingan → makna awal.

Kotak Prompt — “Ringkas Persamaan/Perbedaan + Penyebab Teknis”

Prompt 1 — Ringkas cepat (30 detik):

“Dari tabel komparasi ini, tulis 5 poin: (a) siapa **SOTA**, (b) selisih $\Delta\%$ kami vs **SOTA**, (c) apakah protokol identik, (d) 2 penyebab teknis mengapa hasil bisa beda, (e) 1 risiko bias. Jawab singkat dan tanpa halusinasi.”

Prompt 2 — *Apples-to-apples* check:

“Periksa apakah **dataset**, **metrik**, **protokol** tiap studi sudah identik. Jika tidak, sebutkan *mismatch* dan beri **peringatan** apa yang tidak boleh disimpulkan.”

Prompt 3 — Makna teknis:

“Berdasarkan $\Delta\%$ dan **Catatan**, jelaskan **dua mekanisme teknis** yang paling mungkin menjelaskan selisih, lalu sarankan **1 uji tambahan** untuk konfirmasi.”

Prompt 4 — Versi *camera-ready*:

“Susun paragraf hasil komparasi *camera-ready*: **angka** → **perbandingan** → **makna awal** → **batasan**. Maks 90 kata.”

Ringkasan Super Pendek

- **Bangun tabel:** dataset, metrik/protokol, hasil $\pm$ CI, Δ vs **SOTA**, catatan.
- **Wajib apples-to-apples.** Kalau beda, **beri peringatan**.
- **Tulis narasi 3 kalimat:** **angka** → **perbandingan** → **makna**.
- **Jelaskan penyebab teknis** dan saran uji lanjut.

Dengan ini, perbandingan **objektif**, **tajam**, dan **anti-halusinasi**. Reviewer akan bilang: “clear and fair.”

Menulis Related Work yang “Nendang”

Tugasmu bukan menceritakan semua paper. Tugasmu **memetakan lanskap**, lalu **menunjuk celah** yang **benar-benar kita isi**. Hasil akhirnya: pembaca paham **di mana posisi kita** dan **kenapa kontribusi kita perlu**.

Kata kunci: *cluster tema* → *gap* → **kontribusi unik**.

Langkah 1 — Kumpulkan & Labeli (cepat, rapi, anti pusing)

- Kumpulkan 15–30 paper paling relevan (5 tahun terakhir cukup).
- **Labeli cepat** tiap paper:
 - **Tujuan/masalah**
 - **Metode utama** (*model, rule-based, hybrid*)
 - **Data/domain**
 - **Metrik & protokol**
 - **Kekuatan & limitasi singkat**
- Simpan dalam sheet. Nama kolom sederhana. **Jangan nulis esai dulu**.

Boleh pakai AI untuk *parsing abstrak*, tapi **cek ulang**. Kita anti-halusinasi.

Langkah 2 — Buat *Cluster Tema* (gambar besar dalam 10 menit)

Kelompokkan paper berdasar **benang merah**. Contoh sumbu pengelompokan:

- **Pendekatan/metode**: *CNN, Transformer, rule-based, hybrid*.
- **Jenis data**: teks, gambar, sensor; domain A, B, C.
- **Tujuan spesifik**: deteksi, klasifikasi, prediksi, *retrieval*.
- **Syarat eksperimen**: data kecil vs besar, *few-shot* vs *full-supervised*.

Tips ringkas:

- Buat 3–5 **cluster** saja. **Sedikit, tapi jelas**.
- Beri nama **deskriptif**: “*Transformer untuk domain medis*”, “*Hybrid rules + ML untuk data kecil*”.
- Catat 1–2 **paper jangkar** per cluster (yang paling mewakili).

Kalau mau cepat, boleh minta AI mengusulkan *cluster*. Tetap **validasi manual**.

Langkah 3 — Temukan *Gap* (celah yang *nyata*)

Sekarang cari **ruang kosong** di antara *cluster*.

- **Variasi data** yang belum dicoba (domain baru? *noisy? imbalanced?*).
- **Kondisi uji** yang belum lengkap (tidak ada *robustness test?* tidak lapor *CI/SE?*).

- **Metode** yang belum dikombinasi (belum ada *hybrid* X+Y?).
- **Efisiensi** yang diabaikan (akurasi naik, tapi **latensi** meledak?).
- **Generalitas** lemah (cuma satu dataset? satu bahasa?).

Gap bagus itu: **spesifik, terukur, dan bisa diuji ulang**. Hindari gap yang kabur seperti “belum ada yang sangat optimal”.

Langkah 4 — Kaitkan *Gap* ke Kontribusi Kita

Tunjukkan apa yang kita isi dan **bagaimana** mengisinya:

- **Cluster mana** yang kita teruskan atau gabungkan.
- **Keterbatasan mana** yang kita pecahkan.
- **Eksperimen apa** yang membuktikan isi celah tersebut.

Kalimat sederhana (silakan adaptasi):

- “Berbeda dari *cluster* A yang fokus akurasi, **kami menargetkan efisiensi** pada data skala kecil.”
- “Literatur B belum menguji **domain shift**; kami menguji tiga domain dan melaporkan **CI 95%**.”
- “Tidak ada studi yang menggabungkan X dan Y; **kami usulkan kombinasi X+Y** dan tunjukkan **$\Delta\%$** terhadap *baseline*.”

Struktur Paragraf *Related Work* yang Clean

Gunakan pola 3–4 paragraf pendek. **Kalimat singkat. To the point.**

1. Peta cepat:

“Penelitian terkait terbagi ke **tiga cluster**: (1) *Transformer* untuk domain medis; (2) *hybrid rules+ML* untuk data kecil; (3) *self-supervised* di data tak berlabel.”

2. Ringkas inti per *cluster* + batasan:

“Di *cluster* (1), studi [A,B] kuat di akurasi tetapi **tidak uji robustnes**.

Di (2), studi [C] efisien, **namun hanya satu dataset**.

Di (3), studi [D,E] skala besar, **latensi tinggi**.”

3. Tunjuk *gap* dan kontribusi kita:

“**Celahnya**: uji lintas domain **dengan efisiensi** belum ada.

Kontribusi kami: pendekatan *hybrid* yang **hemat komputasi** dan **stabil** di tiga domain.”

4. Jembatani ke bagian metode/hasil (opsional):

“Bagian berikut menjelaskan rancangan kami dan **bagaimana celah tersebut diuji**.”

Toolkit Mini (boleh jadi catatan pinggir)

- *Mapping*: **cluster** → **gap** → **kontribusi** → **bukti**.
- *Bahasa*: **hindari pujian** (“keren”, “sangat bagus”). Gunakan **fakta**.
- *Istilah asing* ditulis *italic* saat pertama kali muncul.
- **Sitasi hemat tapi tepat**: 2–3 rujukan kunci per *cluster*.

Contoh Kalimat Pembeda Nilai

- “Berbeda dari karya SOTA yang memaksimalkan akurasi, **kami menargetkan robustness** pada **domain baru**.”
- “Tidak seperti studi *few-shot* sebelumnya yang **tanpa CI**, kami melaporkan **CI 95%** dan **n**.”
- “Literatur sebelumnya **belum membahas latensi**, kami mengukur **latensi end-to-end**.”

Kesalahan Umum (hindari)

- Daftar paper berurutan tanpa peta. (**Bikin ngantuk**.)
- Klaim *gap* yang **tidak terukur**.
- Bandingkan *apel* vs *jeruk* (dataset/metrik beda).
- Kutip banyak, **tapi tidak ada benang merah**.

Checklist Siap-Terbit

- Minimal **3 cluster** tema sudah jelas.
- **Gap spesifik** dan bisa diuji ulang.
- **Kontribusi kita** menyasar gap tersebut, bukan yang lain.
- Ada **kata penghubung**: *cluster* → *gap* → kontribusi → bukti.
- **Bahasa ringkas**, angka muncul bila relevan.

Satu Paragraf Contoh (super singkat)

“Literatur terbagi ke **tiga cluster**: *Transformer medis* [A,B] (akurasi tinggi, **tanpa robustnes**), *hybrid rules+ML* [C] (**efisien, satu dataset**), dan *self-supervised* [D,E] (**skala besar, latensi tinggi**). **Celahnya**: belum ada pendekatan **efisien** yang **robust lintas domain**. Kami mengisi celah ini lewat *hybrid X+Y* yang diuji di **tiga domain** dan dilaporkan dengan **CI 95%**.”

Dengan pola ini, bagian *Related Work* kamu **menambah nilai**: bukan ringkasan pasif, tapi **peta strategis** yang menunjukkan **celah nyata** dan **kontribusi asli**. Reviewer suka bagian yang **tajam, jujur, dan fokus**.

Bagian VII — Menentukan Judul Terakhir

Judul yang Meng-highlight Novelty

Judul harus **langsung pamer kontribusi baru**. Bukan puitis. Bukan misterius. **Jelas dan terukur**.

Rumusan cepat: topik – metode – angka/dampak – domain (opsional).

Contoh pola: **Deteksi X – Metode Y – +56% Akurasi – pada Domain Z.**

4 Prinsip Judul yang “Nendang”

- **Sorot *novelty***: apa yang benar-benar baru dari risetmu.
- **Tuliskan angka/dampak** bila ada ($\Delta\%$, *latency*, *throughput*).
- **Sisipkan kata kunci kuat**: “*Akselerasi*”, “*Inovatif*”, “*Baru*”, “*Terobosan*”.
- **Spesifik domain** bila penting: medis, pendidikan, jaringan, manufaktur, dll.

Ingat: judul $\neq$ clickbait. Judul = ringkasan kontribusi.

Template Praktis (tinggal isi)

- [Topik] — [Metode/Arsitektur] — [Angka/Dampak] — [Domain]
- [Aksi/Kata Kuat] [Topik] dengan [Metode] menghasilkan [Angka] pada [Domain]
- [Metode *Inovatif/Baru*] untuk [Topik]: [Angka/Dampak], [Kondisi Uji/Domain]

Kata transisi yang rapi: *dengan, melalui, pada, di, terhadap*.

Kata Kunci Kuat (pilih yang pas)

- **Akselerasi, Inovatif, Baru, Terobosan, Efisien, Ringan, Skalabel, Real-Time, Robust, Hemat Energi, Low-Latency, Interpretable.**

Gunakan 1–2 saja. Jangan menumpuk semua kata kuat.

Contoh Judul (berbagai bidang)

AI/ML

- **Klasifikasi Retinopati — *Transformer* Ringan — +7,9% F1 — pada *Low-Resource* Klinik**
- **Terobosan Deteksi Anomali Sensor — *Hybrid Autoencoder* — -42% *False Alarm***

Jaringan/IoT

- **Akselerasi *Edge Inference* — *Quantized CNN* — 3,1× Lebih Cepat — pada IoT Industri**

- **Routing Inovatif — *Reinforcement Learning* — -28% Latency** di Jaringan Kampus

Sistem/Cloud

- **Penjadwalan Baru Kontainer — *Priority-Aware* — +23% Throughput** di *Multi-Tenant Cluster*

Pendidikan

- **Deteksi Plagiarisme — *Contrastive Embedding* — +12% Precision** — pada Esai Bahasa Indonesia

Kesehatan

- **Screening TB — *Mobile CNN* — Real-Time — -35% Waktu Baca** Radiografi

Checklist 7-Detik (sebelum final)

- [] *Novelty* terlihat tanpa dibaca abstrak?
- [] Ada **angka/dampak** yang bermakna ($\Delta\%$, $\times$, $-ms$, $\downarrow energi$)?
- [] **Metode** disebut jelas (*Transformer, RL, Hybrid*, dll.)
- [] **Topik** spesifik (apa yang dipecahkan).
- [] **Domain** ada jika relevan.
- [] Panjang **ringkas** (12–18 kata ideal).
- [] **Tanpa jargon berlebihan** dan **tanpa klaim bombastis**.

Hindari

- “Studi tentang...” / “Analisis...” **tanpa** hasil atau *novelty*.
- Kata hiperbola kosong: “sangat”, “luar biasa”, **tanpa angka**.
- Akronim internal yang pembaca luar **tidak kenal**.
- Umum banget: “Model Deep Learning untuk Klasifikasi Gambar”.

Cara Menyuntik Angka/Dampak dengan Aman

- Pakai notasi Δ (**delta**) atau tanda $+$ / $-$ / $\times$.
- Cantumkan **metrik** jika muat: *F1, AUC, RMSE, latency, energy*.
- Jika angka **masih berubah**, gunakan istilah dampak **non-angka** yang jelas: *Real-Time, Low-Latency, Hemat Energi*.

Variasi Struktur (opsional, tetap fokus *novelty*)

- **Metode–Topik–Dampak:** *Transformer Medis untuk Segmentasi* — -31% *RMSE*.
- **Dampak–Topik–Metode:** +56% *Akurasi Deteksi Hama dengan CNN Inovatif*.
- **Topik–Dampak–Domain:** *Prediksi Beban* — -24% *MAPE* — pada *PLTS Atap Perkotaan*.

Latihan Kilat (3 langkah)

1. Tulis **Topik** (≤ 5 kata).
2. Pilih **Metode** (1 istilah *teknis*).
3. Tambah **Dampak** (+/-/× + metrik) dan **Domain** (jika perlu).

Gabungkan → **cek dengan checklist** → *done*.

Dengan pola ini, judulmu **tajam, terukur, dan menjual *novelty*** tanpa drama. Reviewer langsung paham **apa yang baru, seberapa besar dampaknya, dan di mana itu berlaku**.

Bagian VIII — Manajemen Referensi & Kepatuhan

Zotero/Mendeley/EndNote: Workflow Anti-Ribet

Kenapa ini penting?

Menulis riset tanpa manajemen referensi itu **capek** dan **rawan salah kutip**.
Pakai *reference manager* agar **rapi, konsisten, dan patuh** ke gaya jurnal.

Target kita: **1 jalur kerja sederhana** dari awal sampai *submit*.

Alur 5 Langkah: Collect → Dedupe → Tag → Cite → Format

Bayangkan seperti *pipeline*. Masuk bersih, keluar kinclong.

1. **Collect (kumpulkan)**
 - Simpan artikel via *browser connector* (Zotero/Mendeley).
 - Ambil **PDF asli + metadata** (judul, penulis, tahun, DOI).
 - Simpan juga **halaman penerbit** untuk verifikasi.
2. **Dedupe (hapus duplikat)**
 - Jalankan fitur **Find Duplicates**.
 - **Satukan** entri kembar. Pilih metadata yang paling lengkap.
 - Hapus file PDF ganda biar hemat ruang.
3. **Tag (beri label)**
 - Buat tag **tema**: *dataset, metode, metrik, domain*.
 - Buat tag **kualitas**: *verified-DOI, no-DOI, predatory-check*.
 - Gunakan **col/collection** per bab agar enak disaring.
4. **Cite (sisipkan sitasi saat menulis)**
 - Pakai *plugin Cite While You Write* (Word/LibreOffice/Docs).
 - Sisipkan sitasi **langsung** dari pustaka.
 - **Jangan ketik manual** (Nama, Tahun)—rawan typo.
5. **Format (gaya jurnal)**
 - Pilih gaya **APA/MLA/Chicago** atau gaya jurnal spesifik.
 - *One-click*: ubah **semua** sitasi dan daftar pustaka otomatis.

Konversi Gaya: APA / MLA / Chicago

- **APA (7th)**: cocok *science* dan *social science*.
- **MLA (9th)**: umum di *humanities*.
- **Chicago**: fleksibel, ada *author–date* dan *notes–bibliography*.
- Tips: simpan **3 profil gaya**. Ubah sesuai jurnal **tanpa edit manual**.

Quality Gate Metadata (wajib lolos)

Sebelum *submit*, cek cepat berikut. Jika 1 gagal, perbaiki.

- DOI valid (bisa *resolve*).
- Judul, penulis, tahun, jurnal/konferensi lengkap.
- Nomor halaman / artikel terisi.
- Penerbit benar (hindari *predatory*).
- PDF sesuai entri (bukan makalah lain).
- Bahasa terdeteksi benar (untuk kapitalisasi judul).
- URL publisher tersimpan (cadangan verifikasi).

Pro tip: buat **saved search** “metadata-missing:true” untuk *to-do list*.

Mini-Checklist “Anti-Ribet”

- [] Semua entri punya DOI/URL.
- [] Tidak ada duplikat.
- [] Minimal 3 tag per entri: *topik, metode, domain*.
- [] Gaya sitasi sama di teks dan pustaka.
- [] Nama penulis konsisten (hindari variasi inisial).
- [] Judul kapitalisasi sesuai gaya (khusus *English title*).

Format organisasi koleksi (opsional, biar enak dicari)

Gunakan pola: **topik–metode–angka/dampak–domain**

- Contoh *collection*:
`DeteksiAnomali–Autoencoder–AUC+0.07–SensorIndustri`
- Contoh **tag**: `topik:deteksi-anomali, metode:autoencoder, dampak:+7% AUC, domain:industri`.

Ini bukan judul paper, ini format penamaan internal agar koleksi gampang dilacak.

Tips Praktis per Alat

Zotero

- Kekuatan: *connector* top-tier, *CSL styles* ribuan, *group library*.
- *Add-on* berguna: **Zotfile** (atur file), **MD Notes** (catatan *markdown*), **DOI Manager**.

Mendeley

- Kekuatan: *PDF annotator* nyaman, *recommendation engine*.
- Pastikan **sinkronisasi** stabil. Sering cek *duplicates*.

EndNote

- Kekuatan: *custom fields*, *journal term lists*.
- Cocok untuk institusi dengan lisensi dan *style* jurnal *proprietary*.

Kotak Prompt: cek konsistensi sitasi & daftar pustaka

Salin ke AI saat draf hampir final.

“Cek konsistensi sitasi dan bibliografi saya.

Input:

1. *In-text citations* (salin 2–3 halaman).
2. Daftar pustaka lengkap.

Tugas:

- Tandai sitasi **ada di teks tapi tidak di pustaka** dan sebaliknya.
- Samakan **format gaya** (sebut gaya target: *APA/MLA/Chicago*).
- Laporkan **duplikat** dan **entri tanpa DOI**.
- Beri saran *auto-fix* di *reference manager* saya.”

Kebiasaan Baik (hemat waktu besar)

- **Cite sambil baca**: saat menemukan paper penting, **langsung simpan, tag, dan anotasi**.
- **Satu referensi = satu catatan ringkas** (3 kalimat: *tujuan, metode, angka kunci*).
- **Rutin dedupe** tiap minggu.
- **Backup** pustaka (ekspor *.bib/ .ris* + folder PDF).

Contoh Mini-Workflow (realistis, 10 menit)

- 0:00–0:03 → Cari 5 paper di *Google Scholar*.
- 0:03–0:06 → **Collect** via *connector* + simpan PDF.
- 0:06–0:08 → **Dedupe** + tambah **tag** (topik, metode, domain).
- 0:08–0:09 → Sisipkan 3 sitasi di draf (**Cite**).
- 0:09–0:10 → Ubah gaya ke **APA (7th) (Format)**. Selesai.

Gotchas (jebakan yang sering bikin kesal)

- **Judul huruf besar semua** di daftar pustaka → perbaiki via **Title Case** sesuai gaya.
- **Nama konferensi** hilang → cek di halaman publisher, isi **booktitle/proceedings**.
- **Akses preprint** tapi sitasi pakai versi jurnal → konsisten pilih salah satu, sebut **DOI** yang benar.
- **Terlalu banyak webpage** → prioritaskan **sumber primer**: jurnal, prosiding, *standard*.

Dengan alur **Collect** → **Dedupe** → **Tag** → **Cite** → **Format**, plus **quality gate metadata**, kamu bisa **patuh gaya**, **anti-halusinasi referensi**, dan **hemat banyak waktu**.

Manajemen referensi bukan kerjaan ekstra. Ini ***engine*** yang bikin naskahmu terlihat profesional.

Transparansi & Replikasi

Kenapa “Transparansi & Replikasi” itu krusial?

Kita ingin riset **bisa dicek** dan **diulang**. Tujuannya sederhana: **percaya karena terbuka**. Ingat dua kata kunci: **transparansi** (*transparency*) dan **replikasi** (*replication*).

Prinsip: **tanpa lampiran data/kode, klaim tinggal cerita**.

Apa yang *idealnya* kamu lampirkan?

Prioritasurut dari paling mudah ke paling kuat.

1. **Lampiran Data (opsional tapi sangat disarankan)**
 - **Data mentah/olah** dalam format terbuka: CSV/Parquet.
 - **Data dictionary** (kamus variabel).
 - **README** cara buka dan pakai datanya.
 - Jika sensitif: **versi *anonymized*** atau **sintetik**.
2. **Lampiran Kode / Notebook**
 - Notebook *Jupyter/RMarkdown* berisi **sel langkah demi langkah**.
 - Skrip **train/eval** yang **bisa jalan ulang**.
 - **requirements.txt** atau *environment.yml*.
 - Opsi lebih rapi: *container (Dockerfile)* untuk *environment* yang sama.
3. **Model & Parameter (jika ada)**
 - *Config file* (hyperparameter, *random seed*).
 - *Checkpoint* model (jika legal dibagi).
 - **Log eksperimen** (mis. *wandb* atau CSV).
4. **Dokumen Bukti**
 - **LOG perubahan** (*changelog*).
 - **Lisensi** file terpisah.
 - **Checklist pelaporan** (lihat di bawah).

Lisensi: pilih yang pas (ringkas dan jelas)

Tujuannya: orang **tahu hak & batas** saat memakai karya kita.

- **Data**: pakai **CC BY 4.0** (atribusi) atau **CC0** (public domain).
- **Kode**: pakai **MIT** (sederhana), **Apache-2.0** (patent-friendly), atau **GPL** (copyleft).
- **Dokumen/gambar**: cocok **CC BY** atau **CC BY-SA**.

Tips cepat

- Tulis lisensi di **LICENSE** (akar repo) + sebut di **README**.
- Kalau campur data+kode, **pisahkan lisensi** masing-masing.

Etika: aman dulu, baru terbuka

Prioritas: lindungi manusia & organisasi.

- **Hapus identitas pribadi (PII)**: nama, NIK, email, GPS.
- Pakai **agregasi / binning** (umur 18–25, dst) bila perlu.
- Minta **izin etik** bila melibatkan manusia/hewan.
- Jika tidak boleh rilis data:
 - Bagikan **statistik ringkas, skema data, kode** yang bisa jalan di **data dummy**, atau **generator data sintetik**.
- Catat **risiko & mitigasi** di README: *bias, limitasi, potensi salah pakai*.

Reproducibility Pack (paket siap ulang)

Bayangkan pembaca ingin menekan **Run** dan selesai. Isinya:

- **README**: tujuan, cara *setup*, cara *run*.
- **Env file**: `requirements.txt` atau `environment.yml`.
- **Run script**: `make train`, `make eval`, atau `bash run.sh`.
- **Data access**: link unduh / instruksi generate / *placeholder*.
- **Config**: `config.yaml` (seed, metrik, *paths*).
- **Notebook**: *EDA* → *training* → *evaluasi* → *grafik*.
- **Lisensi + CITATION.cff** (opsional tapi keren).

Bonus: sediakan **checksum** (mis. SHA256) untuk file besar.

Reporting Checklist (isi sebelum submit)

Wajib isi agar reviewer cepat percaya.

- **Dataset**: sumber, tahun, ukuran (n), etika/izin.
- **Pra-proses**: filter, *imputation*, normalisasi.
- **Protokol uji**: *train/val/test split*, *k-fold*, *seed*.
- **Metrik**: F1/AUC/RMSE/MAE + **CI/SE** jika ada.
- **Baseline**: apa dan kenapa.
- **Perangkat & versi**: OS, CPU/GPU, *library* (versi spesifik).
- **Hasil inti**: angka utama + **CI** + **Δ vs baseline**.
- **Ketersediaan**: link data/kode/lisensi.
- **Batasan**: dimana metode gagal, potensi *bias*.
- **Cara replikasi**: perintah sekali jalan (*one-liner*).

Tangga Transparansi (pilih levelmu)

- **Level 1** — Hanya angka di paper.
- **Level 2** — Angka + kode inti.
- **Level 3** — Angka + kode + *environment* (file *env*).
- **Level 4** — **Siap replikasi penuh**: data (atau pengganti aman), kode, *container*, skrip *one-liner*. 

Target minimal aman: **Level 3**. Target ideal: **Level 4**.

Do & Don't (biar nggak nyasar)

Do

- Tulis **seed** dan **versi library**.
- Kunci konfigurasi di file **.yaml**.
- Sertakan contoh **mini-data** (10–100 baris) untuk cepat cek.

Don't

- Jangan unggah **PII**.
- Jangan ubah angka **setelah** unggah *artifact* (kecuali rilis revisi).
- Jangan simpan **kode rahasia** (API key) di repo publik.

Contoh struktur repo (singkat)

```
project/
├─ README.md
├─ LICENSE
├─ CITATION.cff
├─ data/           # kosong atau link/placeholder
├─ src/            # kode modul
├─ notebooks/     # EDA & evaluasi
├─ configs/        # *.yaml (seed, metrik, path)
├─ scripts/        # run.sh / makefile
├─ env/            # requirements.txt / environment.yml / Dockerfile
└─ results/        # tabel, grafik, log
```

Kotak Prompt (siap pakai ke AI)

1) “Buat paket replikasi minimal”

“Saya punya proyek X. Tolong buat *checklist* dan *template* folder untuk paket **reproducibility**. Sertakan: README langkah jalan, **requirements.txt**, contoh **config.yaml**, dan perintah *one-liner* untuk *train* dan *eval*. Pastikan *seed* dan versi *library* ditulis jelas.”

2) “Pilih lisensi yang cocok”

“Riset saya punya **kode** dan **data**. Kode boleh dipakai ulang dengan atribusi. Data boleh dipakai non-komersial. Rekomendasikan **kombinasi lisensi** yang tepat (mis. MIT untuk kode, CC BY-NC untuk data). Beri teks LICENSE singkat untuk keduanya.”

3) “Audit etika & privasi”

“Tolong *audit* dataset saya untuk risiko **privasi** dan **bias**. Sarankan langkah **anonimisasi**, *aggregation*, atau **data sintetik**. Buat paragraf ‘Pertimbangan Etika & Batasan’ untuk README.”

4) “Ubah notebook jadi skrip siap jalan”

“Konversi notebook saya menjadi skrip `train.py` dan `eval.py`.
Tambah `config.yaml` dan contoh `run.sh` agar hasil utama saya bisa direplikasi dengan satu perintah.”

Intinya

- **Transparansi** = jelas apa yang kamu lakukan.
- **Replikasi** = orang lain bisa mengulang dan mendapat hasil serupa.
- Lampiran **data/kode + lisensi + checklist pelaporan** adalah **tiket ke riset yang dipercaya**.
- Mulai dari kecil. **Naikkan level transparansi** tiap iterasi.

Bagian IX — Siap Kirim & Menjawab Reviewer

Memilih Jurnal & Menyesuaikan Naskah

1) Mindset dulu

Tujuan kita sederhana: **masuk jurnal yang pas**, lalu **lolos desk reject**.
Caranya? **Cocokkan naskah** dengan **aturan jurnal**. Bukan sebaliknya.

Ingat: **kecocokan** mengalahkan **kehebatan**. Salah *fit* = *bye*.

2) Pilih jurnal: 5-langkah kilat

- **Baca aims & scope**. Cek 3–5 artikel terbaru. *Apakah topik itu muncul di sana?*
- **Target pembaca**. Industri? akademik? klinis? Pilih yang **sejalan**.
- **Tipe kontribusi**. Metode baru? *Benchmark*? Aplikasi? Pilih jurnal yang mengundang tipe itu.
- **Kecepatan & etika**. Lihat waktu review rata-rata, *open access* atau tidak.
- **Red flag**. *Scope* terlalu umum, situs berantakan, biaya tidak jelas → **hindari**.

Rule of thumb: jika kamu bisa menulis **1 kalimat** yang menghubungkan karya kamu ke **3 artikel** jurnal itu, **fit-nya kuat**.

3) Cek Aims & Scope (kunci anti desk reject)

Yang harus cocok:

- **Masalah** yang kamu jawab.
- **Metode** utama yang kamu pakai.
- **Domain** data/aplikasi.

Kalimat uji cepat (buat diri sendiri):

“Paper ini **relevan** dengan *aims & scope* karena membahas **[masalah]** dengan **[metode]** pada **[domain]**, yang muncul di artikel **A/B/C** jurnal yang sama.”

Kalau kalimat ini **nggak lancar**, ganti jurnal atau perjelas fokus.

4) Panjang naskah: jangan lewat pagar

Jurnal punya batas. **Ikuti. Titik.**

- **Jumlah kata/halaman**. Biasanya ada *short paper* vs *full paper*.
- **Jumlah gambar/tabel**. Sering dibatasi, mis. **6–8** item.
- **Bahan tambahan**. Taruh sisanya di **Lampiran/Supplement**.

Tips:

- **Padatkan** metode dengan **skema/gambar alur**.

- Simpan tabel detail ke **supplement**.
- **Hapus pengulangan** narasi. Angka cukup sekali, rujuk tabel.

5) Gambar: bersih, tajam, legal

Editor suka grafik yang **jelas**. Reviewer benci yang **buram**.

- **Resolusi**: high DPI. **Label besar, kontras jelas**.
- **Format**: SVG/EPS/PNG (ikuti panduan).
- **CI/SE** kalau komparasi. Tanpa itu, klaim **kurang kuat**.
- **Repro**: sebut **seed**, **n**, dan **protokol** di caption singkat.
- **Hak cipta**: jangan comot gambar orang. Pakai ulang hanya jika **izin/lisensi** jelas.

6) Cover Letter: pintu pertama

Fungsinya: **meletakkan peta** untuk editor. Singkat. Tajam. Ramah.

Struktur 5 kalimat (template cepat):

1. **Hook**: satu kalimat masalah penting di bidangmu.
2. **Novelty**: sebut **3 poin** kontribusi yang **terukur**.
3. **Angka kunci**: “Metode kami **+29% F1** di dataset X; **-18% RMSE** pada Y.”
4. **Fit**: “Karya ini cocok dengan *aims & scope* jurnal karena [alasan jelas] dan merujuk artikel [A,B,C].”
5. **Kepatuhan**: etika, data/kode tersedia, tidak *under review* di tempat lain, tidak ada *conflict*.

Contoh mini (boleh dirombak):

“Dear Editor, kami mengirim naskah berjudul ... yang mengatasi **deteksi anomali waktu nyata**. Kontribusi kami: (1) **metode streaming adaptif**; (2) **dataset industri baru**; (3) **evaluasi terstandar**. Metode kami **meningkatkan F1 29%** pada **Dataset-X** dan **menurunkan RMSE 18%** pada **Dataset-Y**. Topik dan metodenya selaras dengan *aims & scope* jurnal ini serta melanjutkan karya **[Ref A,B,C]**. Semua data/kode kami lampirkan; izin etik telah disetujui.”

7) Highlights: peluru emas

Beberapa jurnal minta **3–5 highlights**. Isinya **klaim terukur**, bukan promosi.

Format efektif:

- **[Masalah] → [Solusi singkat] → [Angka dampak]**
Contoh:
 - “Arsitektur **X-Stream** memproses data sensor **real-time** dengan **latensi <50 ms**.”
 - “Pada **Dataset-A**, model kami **+12,7% AUC** dibanding baseline terbaik.”
 - “Kode & *config* tersedia; **replikasi 1-klik** melalui *container*.”

Hindari:

- Kata umum: “inovatif”, “terobosan”, tanpa angka.
- Klaim besar tanpa **metrik & protokol**.

8) Daftar cek singkat sebelum kirim

- **Fit** dengan **aims & scope** terbukti (tulis kalimat uji).
- **Panjang** sesuai; sisanya di **supplement**.
- **Gambar** tajam; ada **CI/SE** saat komparasi.
- **Cover letter** 5 kalimat; **novelty + angka + fit**.
- **Highlights** 3–5 butir, **kuantitatif**.
- **Etika & lisensi** jelas; **data/kode** tersedia atau *statement* pengganti.

9) Kotak Prompt (siap pakai ke AI)

A. Cocokkan *aims & scope*

“Ini abstrak saya + *aims & scope* jurnal X. Tolong nilai **kecocokan 0–100** dengan alasan 3 poin. Sarankan **2 kalimat** untuk memperkuat *fit* di pendahuluan.”

B. Rapikan *cover letter*

“Ringkas *cover letter* menjadi **5 kalimat**: *hook*, **3 novelty** terukur, **angka kunci**, **fit ke jurnal**, **kepatuhan**. Bahasa formal, padat.”

C. Bikin *highlights* kuantitatif

“Dari tabel hasil ini, tulis **4 highlights** dengan format *Masalah–Solusi–Angka*. Hindari kata promosi.”

D. Pangkas panjang

“Naskah saya kelebihan 1.000 kata. Pangkas pengulangan, pindahkan detail ke *supplement*, dan beri penanda [S1], [S2] di teks.”

Inti yang harus kamu ingat

- **Fit** dulu, hebat kemudian.
- **Angka** mengunci klaim.
- **Cover letter** dan **highlights** adalah **iklan 30 detik** untuk editor.
- **Ikuti aturan** → peluang lolos *desk reject* naik tajam.

Checklist Pra-Submission (Satu Halaman)

Kenapa perlu “satu halaman”?

Agar otak editor **langsung paham**: *apa datamu, apa yang kamu lakukan, hasilnya berapa, dan kenapa penting*. Satu halaman ini jadi **pagar kualitas**. Kalau lulus di sini, **peluang lolos desk reject** naik.

Struktur Satu Halaman (urutan dan isinya)

Bayangkan selembaar kertas A4. **Isi singkat. Padat. Angka dulu**. Setiap baris maksimal 1–2 kalimat. Hindari paragraf panjang.

1. Data (apa & dari mana)

- **Sumber**: publik/privat, izin etik ada.
- **Ukuran**: jumlah sampel/case/jam rekaman.
- **Kualitas**: *missing rate*, outlier, *split train/val/test*.
- **Traceability**: tautkan ke tabel ringkas (*dataset card* mini).

Tulisan contoh: “3 dataset industri (n=18.420), *missing* <2%, *split* 70/15/15, izin etik #123.”

2. Metode (apa yang kamu jalankan)

- **Inti metode**: 1 kalimat.
- **Parameter kunci**: yang paling memengaruhi hasil.
- **Protokol**: *k-fold*, *seed*, jumlah *run*.

Contoh: “Model X-Stream; *window*=128, *lr*=1e-3; 5-fold, 3-run, *seed* ditentukan.”

3. Hasil (angka utama + ketidakpastian)

- **Metrik**: pilih 1–2 utama (mis. F1/AUC/RMSE).
- **Angka**: **mean ± CI/SE**.
- **Kondisi uji**: dataset, *protocol* jelas.

Contoh: “F1=0,842 ±0,02 (5-fold) di Dataset-B; RMSE=0,192 ±0,01 (3-run) di A&B.”

4. ≥3 Novelty (terukur, bukan slogan)

- **Metode baru** → *delta* kinerja/efisiensi.
- **Data/domain baru** → *robustness* terbukti.
- **Analisis baru** → *insight* yang bisa diulang.

Contoh: “(+1) **+12,7% AUC** vs baseline; (+2) **–18% latensi**; (+3) **stabil** saat *domain shift* B.”

5. Abstrak (versi super-ringkas)

- **Gap** → **Tujuan** → **Metode** → **Angka kunci** → **Dampak**.
- 3–4 kalimat cukup.

Contoh: “Ada kekosongan di deteksi anomali *real-time*... Kami mengusulkan... Mencapai **+29% F1**... Memungkinkan pemantauan <50 ms.”

6. Pendahuluan (fit & konteks 3 kalimat)

- **Masalah** di bidangmu.
- **Karya paling dekat** (2–3 studi).
- **Apa yang kamu tambahkan** (singkat, terukur).

Contoh: “Berbeda dari Y dan Z, kami menguji skenario *streaming* dan domain industri.”

7. Referensi (anti-halusinasi)

- **Verifikasi: DOI, penerbit, indeks** (*Scopus/DBLP/PMC*).
- **Sitasi rapi:** konsisten gaya (APA/IEEE).

Catatan: kalau tak ada DOI/publisher jelas → **curigai**.

8. Fig/Table (cukup, tidak berlebihan)

- **Gambar:** ablation bar, *confidence interval*, *delta table*.
- **Tabel:** *Studi–Dataset–Metrik–Nilai–Δ%*.
- **Keterangan:** sebut *n*, *seed*, protokol.

Aturan emas: sedikit tapi tajam. Sisanya di *supplement*.

Gatekeeper: 10 Tes Cepat (centang atau revisi)

- [] **Data jelas:** sumber + *n* + izin etik.
- [] **Protokol fix:** *k-fold/seed/run* tertulis.
- [] **Angka punya CI/SE:** tidak cuma *mean*.
- [] Ada **≥3 novelty** dan semuanya **terukur**.
- [] Abstrak singkat, ada **angka kunci**.
- [] Pendahuluan menyebut **3 studi terdekat**.
- [] Referensi **punya DOI/publisher**; gaya konsisten.
- [] Minimal **1 delta table** (banding baseline).
- [] Minimal **1 ablation bar** (tunjukkan komponen penting).
- [] Semua gambar/tabel **bisa dibaca** (label besar, resolusi tinggi).

Kalimat “uji desk” (self-check 30 detik)

“Dengan dataset **X** (*n*=...) dan protokol **Y**, metode kami memberi **metrik Z** = nilai **±CI**, **unggul Δ%** atas baseline **B**, **stabil** di domain **D**, dan menurunkan **biaya/latensi** sebesar Semua ini **reproducible** dengan kode/data terlampir.”

Jika kalimat di atas belum bisa kamu tulis **dalam sekali tarikan nafas**, perbaiki satu halamanmu.

Mini-Template (boleh copy-paste dan isi angka)

- **Data:** [sumber, n, split, izin]
- **Metode:** [arsitektur/algoritma, parameter kunci, protokol]
- **Hasil:** [metrik1 = mean $\pm$ CI, metrik2 = mean $\pm$ CI]
- **Novelty (≥ 3):**
 1. [$\Delta\%$ performa vs baseline pada dataset ...]
 2. [–% latensi / –% error / +% efisiensi]
 3. [robust di domain ... / insight analitik ...]
- **Abstrak singkat:** [gap $\rightarrow$ tujuan $\rightarrow$ metode $\rightarrow$ angka $\rightarrow$ dampak]
- **Pendahuluan 3 kalimat:** [masalah | 2–3 karya | beda kamu]
- **Referensi:** [DOI valid, gaya konsisten]
- **Fig/Table:** [Delta table #, Ablation bar #, CI/SE ✓]

Kotak Prompt (audit otomatis ke AI)

- **Audit satu halaman**

“Ini *one-page checklist* saya. Tolong cek **kelengkapan & kekuatan** (skor 0–100) per blok: Data, Metode, Hasil, Novelty, Abstrak, Pendahuluan, Referensi, Fig/Table. Kasih **3 saran paling berdampak** + *contoh kalimat* perbaikan.”
- **Perkuat angka & CI**

“Dari tabel hasil ini, hitung **CI 95%** dan **$\Delta\%$ vs baseline**. Tulis ulang bagian ‘Hasil’ dalam **2 kalimat** yang *reproducible*.”
- **Buktikan ≥ 3 novelty**

“Dari ringkasan metode & hasil, rumuskan **3 novelty terukur** (format: *claim* $\rightarrow$ *bukti* $\rightarrow$ *angka* $\rightarrow$ *kondisi uji*).”
- **Bersihkan referensi**

“Ini daftar referensi. **Validasi DOI/publisher**. Tampilkan yang bermasalah dan sarankan pengganti resmi. Rapiakan ke gaya **[APA/IEEE]**.”

Anti-gagal: kesalahan umum (dan obatnya)

- “Hasil kami lebih baik.” **Tanpa angka** $\rightarrow$ **tolak**.
 $\rightarrow$ **Obat:** tulis **metrik + $\pm$ CI + $\Delta\%$ + kondisi uji**.
- Grafik cantik tapi **tanpa CI/SE**.
 $\rightarrow$ **Obat:** tambahkan *error bar*, sebut n, *seed*.
- Referensi **asal mirip judul**.
 $\rightarrow$ **Obat:** cek **DOI & halaman publisher**.
- *Novelty* “inovatif” tapi **tak terukur**.
 $\rightarrow$ **Obat:** ubah ke **target kuantitatif** (waktu, memori, akurasi, stabilitas).

Intinya

Satu halaman = paspor naskahmu.

Kalau **data–metode–hasil–≥3 novelty–abstrak–pendahuluan–referensi–fig/table** sudah **jelas dan terukur**, kamu siap tekan “Submit”.

Strategi Menjawab Reviewer dengan Data

Mindset

- Reviewer bukan musuh. Mereka **sparring partner**.
- Kita menang dengan **data**, bukan emosi.
- Jawaban harus **singkat, spesifik, terukur**, dan **terlacak** ke tabel/fig.

Prinsip emas: “**Claim** → **Bukti** → **Angka** → **Lokasi bukti**.”

Contoh: “Akurasi naik **+6,5%** (Tabel 2, baris 4), $n=1200$, **CI95% $\pm 0,02$** .”

Alur cepat menyiapkan jawaban

1. **Kelompokkan komentar** → *major vs minor*.
2. **Buat rencana uji tambahan** untuk *major*.
3. **Bangun response matrix** (satu baris per komentar).
4. **Tambahkan/ubah** bagian naskah seperlunya.
5. **Tulis jawaban** ringkas, hormat, **berbasis angka**.
6. **Cek silang**: setiap janji di jawaban ada buktinya.

Apa itu *major vs minor*?

- **Major**: menyentuh **validitas utama** (desain eksperimen, baseline salah, *ablation*, *statistical significance*, generalisasi).
- **Minor**: **perapihan** (typo, gaya sitasi, label sumbu, kalimat kurang jelas).

Taktik:

- *Major* → **uji/analisis tambahan** + perbaikan teks.
- *Minor* → **perbaiki langsung** + sebut lokasi perubahan.

Paket “Tambahan Uji/Analisis” yang cepat & berdampak

Pilih 2–4 yang paling relevan. Jangan semua sekaligus.

- **Ablation singkat**: matikan modul inti → ukur $\Delta\%$.
- **Robustness check**: tambah satu **domain/dataset baru** kecil (n cukup) → laporkan **stabil/turun**.
- **Sensitivity analysis**: geser **hyperparameter kunci** ($\pm 20\%$) → lihat tren.
- **Significance test**: hitung **CI95%** dan **p-value** → tampilkan di *delta table*.
- **Error analysis**: 10–20 **contoh salah klasifikasi** → pola kesalahan + solusi singkat.
- **Compute/latency audit**: laporkan **waktu, memori, FLOPs** → bukti efisiensi.
- **Data leakage check**: buktikan **split** bersih (ID unik, *no overlap*).

Tips: simpan hasil tambahan sebagai **Gambar S1–S3** atau **Tabel S1–S2** di *supplement*. Naskah utama tetap ramping.

Response Matrix (template siap pakai)

Salin ke Google Docs/Sheets. Satu baris = satu komentar.

#	Ringkasan Komentar	Jenis	Respons Inti (1–2 kalimat)	Bukti & Angka	Perubahan di Naskah
1	Baseline kurang kuat	Major	Kami menambah baseline Z yang direkomendasikan. Metode kami tetap unggul.	F1 +4,2% vs Z, CI95% $\pm 0,015$, $n=1200$ (Tabel S1).	Hal. 7, Bag. 4.2; Tabel S1 baru.
2	Perlu <i>ablation</i> modul X	Major	Kami jalankan <i>ablation</i> X. Kinerja turun, artinya X penting.	ΔRMSE -10,3% saat X dimatikan (Gambar S2).	Hal. 9, paragraf 2; Gambar S2.
3	Label sumbu kurang jelas	Minor	Kami perjelas label & unit.	Figur 3 kini menyertakan “ms” dan “MB”.	Hal. 11; Figur 3 diperbarui.

Kunci: kolom “Bukti & Angka” **wajib** menyebut **metrik, Δ , CI/SE, n, lokasi**.

Contoh jawaban *major* (padat & data-driven)

Komentar R1-M2: “Bandingkan dengan model Z dan laporkan signifikansi.”

Jawaban:

“Terima kasih. Kami menambahkan perbandingan dengan **model Z** sesuai saran. Pada Dataset-B, metode kami mencapai **F1=0,842 $\pm$ 0,02** (5-fold), sedangkan Z **0,804 $\pm$ 0,02**, selisih **+4,7%**. Uji *paired t* menunjukkan **p=0,008**. Rinciannya ada di **Tabel S1**, dan kami menyisipkan ringkasan pada **Hal. 7, Bag. 4.2**.”

Contoh jawaban *minor*

Komentar R2-m3: “Judul Gambar 3 kurang informatif.”

Jawaban:

“Kami telah memperbarui judul **Gambar 3** dan menambahkan unit (**ms**) serta **CI95%** pada *error bar*. Lihat **Hal. 11**.”

Pola kalimat yang aman dipakai

- “Kami **setuju** dan telah ...” (jika sejalan).
- “Kami **menghargai** masukan ini. Untuk menguji hal tersebut, kami ...” (netral).
- “Kami **klarifikasi**: ...” (jika ada salah paham, sertakan angka).
- “Kami **tambahkan analisis** pada **Tabel/Gambar S...**” (arahkan ke bukti).

Hindari: debat panjang tanpa data, kata sifat kosong (*keren, unggul, mutakhir*).

Mini-Playbook memilih eksperimen tambahan (≤ 1 minggu kerja)

- Dampak ke klaim utama tinggi? Ambil.
- Biaya komputasi masuk akal? Ambil.
- Menjawab langsung pertanyaan reviewer? Ambil.
- Redundan? Buang.

Prioritas S-tier: **baseline kuat, ablation inti, significance & CI, robustness domain**.

Kotak Prompt — draf jawaban *major/minor comments*

1) Ringkas & klasifikasi komentar

“Ringkas semua komentar di bawah ini. Kelompokkan jadi **major** dan **minor**. Beri **alasan** klasifikasi (1 kalimat per komentar).”

2) Rekomendasi uji tambahan (terukur)

“Untuk komentar **major**, usulkan **3 uji/analisis** paling berdampak. Untuk tiap uji, spesifikkan **metrik, n, protokol, estimasi CI**, dan **output** yang akan jadi **Tabel/Gambar S#**.”

3) Tulis draf jawaban *major*

“Tulis **jawaban singkat** (maks 5 kalimat) untuk komentar **[paste komentar]**. Ikuti pola **Claim** → **Bukti** → **Angka** → **Lokasi**. Gunakan istilah *statistik* yang tepat (*CI/95%, p-value, effect size*).”

4) Tulis draf jawaban *minor*

“Tulis jawaban **1–2 kalimat** untuk minor ini: **[paste komentar]**. Sertakan **lokasi halaman/figur** yang diubah.”

5) Bangun *response matrix*

“Buat tabel **response matrix** dengan kolom: # | Ringkasan Komentar | Jenis | Respons Inti | Bukti & Angka | Perubahan di Naskah. Isi dari daftar komentar & hasil uji terbaru saya.”

Quality Gate sebelum kirim balik

- [] Semua *major* ada uji/analisis atau klarifikasi angka.
- [] Setiap jawaban menunjuk ke **Tabel/Gambar/Halaman** yang spesifik.
- [] Angka pakai **metrik + $\Delta\%$ + CI/SE + n**.
- [] Nada **hormat** dan **konsisten** (tanpa defensif).
- [] *Supplement* rapi: **Tabel/Gambar S#** tersusun & dirujuk.

Intinya

Menang melawan reviewer = menang dengan data. Tambahkan uji seperlunya, tulis **respons terukur**, dan kunci dengan **response matrix** yang jelas. Selesai? **Submit revisi.** Lanjut ke *accept*.

Lampiran A: Template Tabel Komparasi Kuantitatif (siap isi)

Kenapa tabel komparasi itu penting

- **Membuktikan klaim** kamu dengan angka yang jelas.
- **Apples-to-apples**: dataset sama, *protocol* sama, metrik sama.
- **Mudah dibaca reviewer**. 10 detik paham inti hasil.

Ingat: **klaim tanpa angka = opini**. Klaim dengan angka + konteks uji = **ilmu**.

Prinsip cepat (ingat terus)

- **Satu metrik utama** per tabel (mis. akurasi, F1, RMSE).
- **Satu protokol uji** per baris (jelas *split*, *fold*, *seed*).
- Tulis **Δ absolut** dan **$\Delta\%$** vs *baseline* atau *SOTA*.
- Sertakan **ketidakpastian**: *SD/SE* atau *CI*.
- Catat **n**, *hardware*, **waktu**, **memori** bila relevan.

Kolom wajib (ringkas tapi padat)

- **Studi (Tahun)**
- **Dataset / Domain**
- **Metrik** (jelas: akurasi %, F1, RMSE, AUC, dsb.)
- **Protokol** (5-fold, holdout 20%, seed, preprocessing)
- **Nilai ($\pm$ SD/CI)**
- **Δ vs Baseline / SOTA** (abs & %)
- **n** (jumlah sampel/run)
- (opsional tapi keren): **Waktu (s)**, **Memori (MB)**, **Hardware**

Template tabel — siap isi

Salin–tempel, lalu ganti isinya. Jaga konsistensi satuan & metrik, ya.

Studi (Tahun)	Dataset/ Domain	Metrik (unit)	Proto kol Uji	Nilai ($\pm$ SD/ CI)	Δ vs Basel ine (abs)	Δ (%)	n	Wak tu (s)	Mem ori (MB)	Hardw are
Paper- A (2023)	X (split resmi)	Akura si (%)	5-fold, seed= 42	88.2 $\pm$ 0.9	—	—	12 00	52	380	RTX 2060

Paper-B (2022)	X (varian B)	Akura si (%)	5-fold, seed= 42	89.1 ± 0.8	+0.9	+1.0 %	90 0	55	410	RTX 2070
SOTA (2024)	X (noisy)	Akura si (%)	5-fold, seed= 42	90.0 ± 1.0	+1.8	+2.0 %	12 00	61	420	A100 (40GB)
Kita (2025)	X (split resmi)	Akura si (%)	5-fold, seed= 42, aug= Y	91.4 ± 0.7	+3.2	+3.6 %	12 00	58	350	RTX 3060 (12GB)

Catatan: semua uji memakai *preprocessing* yang sama; *learning-rate*, *batch*, dan *seed* konsisten.

Cara isi (langkah super singkat)

1. **Pilih satu metrik utama.** Contoh: **Akurasi (%)**.
2. **Samakan protokol.** *5-fold*, *seed*, *split*, *augmentation* → harus konsisten.
3. **Ambil angka dari paper.** Tulis **Nilai (±SD/CI)**, sebutkan sumber.
4. **Hitung Δ:**
 - $\Delta \text{ absolut} = \text{nilai_kita} - \text{nilai_pembanding}$.
 - $\Delta\% = ((\text{nilai_kita} - \text{nilai_pembanding}) / \text{nilai_pembanding}) \times 100\%$.
5. **Tambahkan n, waktu, memori, hardware.** Ini bikin klaimmu **lebih kredibel**.
6. **Cek satuan.** Jangan campur **detik** dan **ms** di kolom yang sama.
7. **Beri catatan** bila ada perbedaan kecil (mis. varian dataset).

Checklist 10 detik sebelum submit

- [] Dataset & *protocol* sama untuk semua baris.
- [] Metrik konsisten dan **jelas satuannya**.
- [] Ada **SD/SE/CI** atau minimal keterangan variasi.
- [] **Δ abs** dan **Δ%** tertulis rapi.
- [] **n, waktu, memori, hardware** diisi bila relevan.
- [] Tidak ada *cherry-picking* metrik.

Kesalahan umum (hindari!)

- “A lebih baik dari B” **tanpa angka**.
- Bandingkan *protocol* berbeda, lalu klaim **menang**.
- Pindah-pindah metrik demi terlihat unggul.

- Lupa tulis **n** dan **ketidakpastian** (CI/SD).
- Campur dataset berbeda **tanpa** catatan.

Kalimat siap pakai (isi angkanya)

- “Model kami mencapai **91,4% $\pm$ 0,7** pada Dataset X (5-fold, seed=42), **+3,2 pts (+3,6%)** terhadap **SOTA (2024)**.”
- “Waktu inferensi turun **15% (61 s $\rightarrow$ 58 s)** dengan memori **–16%**, akurasi tetap dalam **CI 95%** yang tumpang tindih.”
- “Keuntungan terbesar terlihat pada subset noisy (**+4,5 pts**), konsisten di 3 run (**$SD \leq 0,8$**).”

Kotak Prompt (biar cepat nulisnya)

- **Ringkas tabel $\rightarrow$ paragraf 3 kalimat**
“Dari tabel ini, tulis 3 kalimat: (1) angka inti terbaik, (2) Δ absolut & $\Delta\%$, (3) makna praktisnya. Hindari opini.”
- **Cek fairness**
“Apakah protokol uji, split, dan seed sudah identik? Jika tidak, beri catatan koreksi.”
- **Stabilitas hasil**
“Tunjukkan SD/CI tiap studi dan simpulkan apakah perbedaan praktis berarti.”
- **Efisiensi**
“Ringkas perbandingan waktu & memori. Apakah ada trade-off signifikan?”

Mini tantangan (15 menit)

- Isi template dengan **3 studi + hasilmu**.
- Hitung **Δ absolut & $\Delta\%$** .
- Tulis **1 paragraf**: angka $\rightarrow$ perbandingan $\rightarrow$ makna.
- Bonus: tambah baris **waktu** dan **memori**.
- Cek lagi **apples-to-apples**. Kalau belum, perbaiki *protocol*.

Intinya

- **Tabel komparasi** adalah senjata anti-halusinasi.
- **Samakan protokol**.
- **Laporkan ketidakpastian**.
- **Tulis Δ dengan jujur**.
- Biar data yang ngomong. Kamu tinggal menyimpulkan.

Lampiran B: Kamus Metrik & Rumus (efek, CI, uji cepat)

Kenapa kita butuh “Kamus Metrik & Rumus”?

Banyak mahasiswa bisa menulis paragraf, tapi **tidak bisa menjelaskan angka**. Padahal dalam penelitian, **angka adalah bahasa universal ilmu pengetahuan**.

Tanpa paham metrik, kamu seperti **bermain bola tanpa tahu skor**. Kamu berlari, tapi tidak tahu menang atau kalah. Jadi, *Lampiran B* ini adalah “kamus kecil” yang membantu kamu **ngerti arti angka** — cepat, jelas, dan tidak ngawang.

Tujuan utamanya:

- Biar kamu bisa menjelaskan hasil riset dengan percaya diri.
- Biar reviewer kagum karena kamu pakai metrik dan rumus yang benar.
- Biar AI yang bantu kamu nulis juga ngerti konteks metriknya (anti-halusinasi).

Jenis-jenis metrik yang wajib kamu kenal

Berikut daftar *metrik* paling umum yang sering muncul di jurnal:

1. Metrik performa model (Machine Learning / Statistik)

- **Accuracy (%)**: seberapa banyak prediksi benar.
→ “Model A punya akurasi 92%, berarti 92 dari 100 prediksi benar.”
- **Precision & Recall**: *Precision* = ketepatan; *Recall* = seberapa banyak yang berhasil ditemukan.
→ Gunakan bareng di **F1-score = $2 \times (\text{Precision} \times \text{Recall}) / (\text{Precision} + \text{Recall})$** .
- **RMSE (Root Mean Squared Error)**: seberapa besar rata-rata kesalahan prediksi.
→ Semakin kecil, semakin baik.
- **AUC (Area Under Curve)**: kemampuan model membedakan kelas positif vs negatif.
→ Semakin mendekati 1, semakin jago modelnya.

2. Metrik efisiensi sistem

- **Latency (ms)**: waktu yang dibutuhkan sistem untuk merespons.
→ Idealnya cepat, tapi tetap akurat.
- **Throughput**: berapa banyak data yang bisa diproses per detik.
- **Memory Usage / GPU Time**: tunjukkan efisiensi komputasi.

3. Metrik validasi eksperimen

- **n**: jumlah *sample* atau *run* (semakin besar, semakin stabil hasilnya).
- **SD (Standard Deviation)**: seberapa menyebar hasil uji dari rata-ratanya.
- **CI (Confidence Interval)**: kisaran angka yang menunjukkan *seberapa yakin* kita terhadap hasil.
→ Contoh: “91.4% ± 0.7” artinya kemungkinan besar (95%) nilai sebenarnya ada di antara 90.7–92.1%.

Rumus-rumus cepat yang wajib diingat

Supaya kamu bisa langsung konversi klaim ke angka.

Nama Rumus	Fungsi	Rumus Matematis	Contoh Praktis
Relative Improvement	Mengukur peningkatan performa relatif	$((A - B) / B) \times 100\%$	Model baru naik 8% dibanding baseline
Error Reduction	Mengukur seberapa banyak error berkurang	$((E_B - E_A) / E_B) \times 100\%$	Error turun 30% dari model lama
Effect Size (Cohen's d)	Mengukur besar efek perbedaan	$(\text{mean}_1 - \text{mean}_2) / \text{SD}_{\text{pooled}}$	Nilai 0.8 = efek besar
Confidence Interval (95%)	Batas kepercayaan hasil	$\text{mean} \pm 1.96 \times (\text{SD} / \sqrt{n})$	$90 \pm 1.5 \rightarrow \text{CI: } [88.5, 91.5]$

Tips cepat mengingat konsep penting

CI = percaya diri berbasis data.

Kalau CI kecil → hasil stabil.

Kalau CI besar → masih goyang, perlu data tambahan.

Effect size itu kayak “*seberapa berartinya perbedaanmu*”.

Kadang beda 1% akurasi, tapi kalau data besar, efeknya bisa sangat signifikan.

n penting banget.

Tanpa *n*, angka kamu bisa jadi ilusi. Coba bayangin bilang “naik 10%” padahal cuma dari 5 sampel.

Mini *Playbook* uji cepat (Quick-Check)

Kamu bisa cek hasil eksperimen hanya dengan beberapa baris *Python* atau bahkan di kalkulator.

Cek peningkatan performa:

$$\text{delta} = ((A - B) / B) * 100$$

Contoh: A=0.91, B=0.85 $\rightarrow \Delta = 7.05\%$

Hitung CI sederhana:

$$CI_{95} = \text{mean} \pm 1.96 * (\text{sd} / \text{sqrt}(n))$$

Contoh: mean=91.4, sd=0.7, n=5 $\rightarrow CI = [91.1, 91.7]$

Bandingkan dua model:

```
diff = abs(A - B)
if diff < CI: hasilnya belum signifikan
```

$\rightarrow$ Simpel, tapi cukup buat tahu apakah peningkatan nyata atau hanya kebetulan.

Untuk yang suka tantangan:

Coba buat “mini kamus metrik” kamu sendiri dari eksperimen di lab:

- Tulis **nama metrik**, **fungsi**, dan **rumusnya**.
- Tambahkan **contoh angka dari eksperimenmu**.
- Tulis **interpretasi**: “Semakin kecil $\rightarrow$ lebih baik” atau “Semakin besar $\rightarrow$ lebih baik”.
- Simpan di *spreadsheet* atau *Google Sheet* supaya bisa auto-calc.

Bonus: buat AI kamu (ChatGPT, Copilot, atau Ollama) bantu mengecek metrik itu — biar setiap kali kamu tulis “Model A lebih baik”, AI otomatis tanya:

“Berapa persen lebih baiknya?”

Inti

- Angka bukan hiasan, tapi bukti ilmiah.
- Metrik adalah bahasa yang harus kamu kuasai sebelum bicara “lebih baik”.
- Rumus-rumus di sini bukan matematika rumit — tapi kompas untuk tahu arah risetmu.

Dan yang paling penting:

Kuasai metrik → kuasai narasi risetmu.

Karena di dunia sains, *data talks — bullshit walks*.

Lampiran C: Kumpulan Prompt Siap Pakai

Kenapa perlu *kumpulan prompt*?

Menulis dengan AI itu seperti naik sepeda listrik. **Kamu tetap mengayuh**, AI hanya **mengakselerasi**. Dengan *prompt* yang tepat, kamu dapat:

- **Hemat waktu** (AI langsung ke inti).
- **Anti-halusinasi** (AI selalu minta bukti).
- **Konsisten format** (tabel, angka, dan istilah rapi).

Cara pakai: ganti teks di **<ANGKA>**, **<TABEL_X>**, **<DATASET>**, dll. Tambah konteks jika perlu. Simpan *prompt* favoritmu di *notes*.

0) *Meta-prompt* (pengaman anti-halusinasi)

Tujuan: memasang *guardrail* sebelum bekerja.

Prompt:

“Bertindak sebagai *asisten penulisan ilmiah* yang **data-first** dan **anti-halusinasi**.

- Jika data kurang, **tanyakan**.
- Setiap klaim komparatif harus ada **metrik, protokol, n, CI/SD**.
- Jika menyebut studi, sebutkan **DOI/link penerbit**.
- Output selalu beri **sumber/jejak data** (*audit trail*).
Konfirmasi bahwa Anda paham aturan ini.”

1) *EDA cepat & tajam* (dari tabel/grafik)

Tujuan: menemukan pola/anomali paling penting.

Prompt:

“Dari **<TABEL_X>** (kolom: **<DAFTAR_KOLOM>**), ringkas **3 pola** dan **1 anomali** paling signifikan.

- Jelaskan dengan angka (mean/median/ $\Delta\%$).
- Beri dugaan penyebab (maks 1 kalimat/poin).
- Akhiri dengan **saran uji lanjut** (2 butir).
Jangan membuat data baru. **Jika ragu, tulis ‘perlu data tambahan’.**”

2) Menulis metode yang *reproducible*

Tujuan: narasi metode dari daftar parameter.

Prompt:

“Susun narasi **Metode** dari daftar berikut:

Alat: <SOFTWARE/VERSI>, Model: <MODEL>, *learning-rate* <LR>, batch <BATCH>, *seed* <SEED>, *augmentation* <AUG>.

- Tulis ringkas, urut **pipeline**.
- Cantumkan **protokol uji** (fold, split, *seed*).
- Tambahkan **tabel parameter** (nama, nilai).
- Tutup dengan kalimat **reproducibility** (versi, *commit hash*, *hardware*).”

3) Menulis hasil: angka dulu, tafsir sesudah

Tujuan: paragraf hasil yang rapi dan tajam.

Prompt:

“Buat paragraf **Hasil** dari <TABEL_SKOR>:

- Kalimat 1: **angka inti** (mean ± SD/CI, n).
- Kalimat 2: **perbandingan** vs *baseline* dan $\Delta\%$.
- Kalimat 3: **makna awal** (singkat, tanpa spekulasi).
Larang kata ‘lebih baik’ tanpa angka. **Tunjukkan unit metrik.**”

4) Mengkuantifikasi klaim (wajib angka)

Tujuan: ubah ‘A lebih baik’ → ‘A **x%** lebih baik’.

Prompt:

“Ubah kalimat kualitatif berikut menjadi **pernyataan kuantitatif** memakai rumus:

- *Relative improvement* = $((A-B)/B) \times 100\%$
 - *Error reduction* = $((E_B - E_A)/E_B) \times 100\%$
- Input: A=<A>, B=, E_A=<EA>, E_B=<EB>, n=<N>, CI=<CI>.
Hasilkan kalimat: ‘**A** = < Δ_{abs} > (< Δ_{rel} >%) vs B, n=, CI=’.
Jika data kurang, tulis **variabel apa yang kurang.**”

5) Visual kuantitatif yang meyakinkan

Tujuan: *brief* figur untuk peneliti/grafis.

Prompt:

“Rancang 1–2 **usulan grafik** dari **<TABEL_HASIL>**:

- Grafik 1: **bar ± CI/SE** untuk metrik utama.
- Grafik 2: **ablation bars** (fitur → Δ).
Jelaskan **judul figur**, **sumbu**, **catatan protokol**, dan **pesan kunci** dalam 3 poin. Tanpa *chartjunk*.”

6) Cek *novelty* cepat (5 tahun terakhir)

Tujuan: memetakan posisi karya.

Prompt:

“Bandingkan temuan kita dengan **3–5 studi** paling relevan 5 tahun terakhir (sebut **DOI**).

Buat ringkasan: **dataset**, **metode**, **metrik**, **hasil**, **protokol**.

Tutup dengan **tabel Δ vs SOTA** (apples-to-apples).

Jika referensi tidak terverifikasi, beri label ‘**butuh verifikasi DOI**’.”

7) *Tightening* abstrak (150–250 kata)

Tujuan: ringkas, fokus *novelty*.

Prompt:

“Perbaiki **Abstrak** (150–250 kata) dengan struktur: **gap** → **tujuan** → **metode** → **angka kunci** → **dampak**.

- Sorot **3 angka** paling penting (dengan unit).
- Hindari klaim luas tanpa angka.
- Beri **1 kalimat batasan** dan **1 kalimat dampak**.
Outputkan **2 versi**: *fokus aplikasi* dan *kontribusi teoretis*.”

8) Pendahuluan tanpa halusinasi

Tujuan: jahit *background* → *research gap* → *objectives* berbasis data.

Prompt:

“Tulis **Pendahuluan** ringkas yang:

- (1) menyajikan **latar** berbasis angka,
- (2) menegaskan **research gap** yang spesifik,
- (3) menyatakan **objectives** yang terukur.

Cantumkan 4–6 referensi dengan **DOI/publisher page**.

Jika ada yang meragukan, beri tag **[butuh verifikasi]**.”

9) *Related Work* yang adil & *apples-to-apples*

Tujuan: ringkas persamaan/perbedaan + penyebab teknisnya.

Prompt:

“Ringkas **related work** dalam 5–7 poin:

- **Persamaan** metodologis vs kita.
- **Perbedaan** protokol/dataset/fitur.
- **Penyebab teknis** kenapa hasil berbeda (hipotesis singkat).
Akhiri dengan 1 **tabel komparasi**: *studi* | *dataset* | *metrik* | *protokol* | *hasil* | Δ vs *kita*. Tanpa opini berlebihan.”

10) Tabel komparasi kuantitatif (siap isi)

Tujuan: tabel yang konsisten dan mudah salin ke *Google Docs*.

Prompt:

“Bangun tabel komparasi dengan kolom:

Studi (Tahun) | **Dataset/Domain** | **Metrik (unit)** | **Protokol Uji** | **Nilai ($\pm$ SD/CI)** | **Δ vs Baseline (abs | %)** | **n** | **Waktu (s)** | **Memori (MB)** | **Hardware**.

Isi dari `<TABEL_SUMBER>`.

Validasi: metrik sama, protokol sama, sebut **seed** bila ada. Tambah **catatan footer** jika ada variasi.”

11) *Response to reviewers* (pakai data)

Tujuan: jawab *major/minor comments* secara elegan dan terbukti.

Prompt:

“Buat *response matrix* dari `<DAFTAR_KOMENTAR>`:

- **Ringkas komentar** (1 kalimat).
- **Respons inti** (1–2 kalimat, sopan).
- **Bukti/data**: angka baru/figur tambahan/CI/uji sensitivitas.
- **Status**: *resolved/partially/plan for revision*.
Jika diminta uji tambahan, usulkan **eksperimen minimal tetapi kuat** (n, metrik, protokol).”

12) *Quality Gate* sitasi & referensi

Tujuan: memastikan daftar pustaka rapi dan valid.

Prompt:

“Audit **sitasi** dan **daftar pustaka** pada naskah:

- Cek **konsistensi gaya** `<APA/MLA/Chicago>`.

- Verifikasi **DOI** dan **publisher page**.
- Tandai **duplikasi** dan **metadata hilang**.
- Hasilkan **daftar koreksi** + *BibTeX* bersih.”

13) *One-liner prompts* (super singkat, super nendang)

- “Ubah klaim ini jadi angka (sertakan $\Delta\%$ dan CI): **<KALIMAT>**.”
- “Deteksi tiga pola + satu anomali dari **<TABEL_X>**.”
- “Kompres abstrak ke 180 kata, jaga 3 angka kunci.”
- “Usulkan figur bar $\pm$ CI untuk **<METRIK>** dari **<TABEL_Y>**.”
- “Bandingkan vs **SOTA** secara *apples-to-apples* (metrik & protokol sama).”

Penutup

Intinya: *prompt* yang tajam = tulisan yang tajam. Selalu sertakan **data**, **protokol**, **n**, **CI/SD**, dan **DOI**. Dengan paket *prompt* di atas, kamu bisa:

- menulis lebih cepat,
- lebih **akurat**,
- dan **disukai reviewer**.

Gaskeun. Data jalan dulu, narasi menyusul.

Glossary

A

- **Abstrak**
Ringkasan paper: **gap** → **tujuan** → **metode** → **angka kunci** → **dampak**. Panjang 150–250 kata. Hindari klaim tanpa angka.
- **Ablation** (*ablation study*)
Uji melepas satu komponen/fitur untuk lihat **dampaknya**.
Tujuan: tahu bagian mana yang **benar-benar kontribusi**.
Contoh: tanpa *augmentation*, akurasi turun **−2.4%**.
- **Accuracy** (*akurasi*)
Proporsi prediksi benar. Cocok untuk kelas **seimbang**.
Jangan pakai sendirian saat kelas **imbalance**.
- **Apples-to-apples**
Perbandingan **adil**: metrik, dataset, **protokol uji sama**.
Larangan: bandingkan hasil *paper* lain dengan protokol berbeda.
- **Audit Trail**
Jejak bukti dari klaim → tabel → data mentah.
Manfaat: **anti-halusinasi**, mudah direplikasi.

B

- **Baseline**
Titik awal pembandingan. Bisa model lama atau metode standar.
Tip: selalu laporkan **Δ vs baseline** (selisih absolut & %).
- **Batch Size**
Jumlah sampel per langkah training. Pengaruh ke **stabilitas & memori**.
- **Bias**
Kecenderungan sistematis yang menggeser hasil. Kurangi dengan desain uji yang adil.

C

- **CI** (*confidence interval*)
Rentang keyakinan nilai sebenarnya. Tulis format: **mean ± CI95%**.
Contoh: 91.4 ± 0.7 (CI95%, n=1200).
- **Cross-Validation** (*k-fold CV*)
Data dibagi *k* lipatan. Latih & uji bergiliran.
Kelebihan: estimasi performa **lebih stabil**.

D

- **Data-First**

Menulis **setelah** ada tabel/grafik/angka. Narasi mengikuti data.

Mantra: “Tanpa data, jangan menulis.”

- **Δ (Delta)**

Selisih hasil.

- **Δ_abs:** perbedaan angka langsung (mis. +3.2 poin).
- **Δ_%:** perbedaan relatif (mis. +3.6%).

E

- **EDA** (*exploratory data analysis*)

Cek distribusi, tren, korelasi, anomali.

Output cepat: 3 pola + 1 anomali + saran uji.

- **Effect Size**

Ukuran **kebesaran efek**, bukan hanya *significance*.

Contoh: *Cohen's d* kecil (0.2), sedang (0.5), besar (0.8).

- **Error Reduction**

Penurunan error relatif: $((E_B - E_A) / E_B \times 100\%)$.

Cocok saat metrik utama berbasis **error** (mis. WER, RMSE).

F

- **F1-Score**

Rata-rata harmonik *precision* dan *recall*.

Pakai saat kelas **imbalance**.

G

- **Guardrail**

Aturan *prompt* supaya AI **tidak halus**. Contoh: “Semua klaim wajib metrik, n, CI, dan protokol.”

H

- **Halusinasi (AI)**

AI menjawab **tanpa bukti**.

Obatnya: verifikasi DOI, kutip tabel/angka, simpan *provenance*.

- **Hardware**

Spesifikasi komputasi. Cantumkan tipe GPU/CPU, RAM, VRAM.

Kenapa penting: hasil bisa **berubah** jika perangkat beda.

I

- **Imbalance**
Kelas tidak seimbang.
Solusi: metrik **F1/AUC**, *stratified split*, *reweighting*.

L

- **Learning Rate**
Langkah pembaruan bobot saat training. Terlalu besar → **diverge**; terlalu kecil → **lambat/underfit**.

M

- **MAE / RMSE**
 - **MAE:** rata-rata **selisih absolut**. Mudah dibaca.
 - **RMSE:** penalti lebih besar untuk **error besar**.
- **Memori (MB)**
Jejak pemakaian RAM/VRAM saat latihan atau inferensi.
Lapor bersama waktu (detik).
- **Metrik (unit)**
Ukuran kinerja + **unit**. Contoh: akurasi (%) / waktu (s).
WAJIB: sebut **unit** dan **konteks uji**.

N

- **Novelty**
Kebaruan yang **terukur**: metode, data/domain, performa, efisiensi, interpretasi.
Target: minimal **3** novelty.
- **n (sample size)**
Jumlah contoh uji. Tulis jujur. Pengaruh ke **stabilitas & CI**.

O

- **Objectives**
Tujuan riset yang **spesifik** dan **terukur**.
Contoh: “Menaikkan F1 ≥ **+3%** pada dataset X, protokol sama.”
- **Overfitting**
Model jago di train, **jelek di test**.
Deteksi: selisih performa train vs valid.
Obat: regularisasi, data lebih, *early stopping*.

P

- **p-value**
Probabilitas hasil minimal setajam itu jika **nol hipotesis benar**.
Ingat: kecil $\neq$ efek besar. Laporkan **effect size** juga.
- **Protokol Uji**
Aturan evaluasi: split, *seed*, *k-fold*, preprocessing.
WAJIB konsisten untuk *apples-to-apples*.
- **Provenance**
Asal-usul data/angka/figur. File apa, *commit* berapa, siapa jalankan.

R

- **Reproducibility**
Orang lain bisa **mengulang** hasil yang sama.
Checklist: seed, versi alat, *commit hash*, skrip, *requirements.txt*.
- **Relative Improvement**
Kenaikan relatif: $((A-B)/B \times 100\%)$.
Contoh kalimat: "Metode kami **+3.6%** vs baseline (seed=42, 5-fold)."

S

- **Seed**
Angka acak untuk **konsistensi**. Selalu laporkan.
Contoh: seed=42.
- **Sensitivity Analysis**
Uji seberapa sensitif hasil terhadap perubahan kecil parameter.
Tujuan: cek **robustness**.
- **SOTA (*state of the art*)**
Performa terbaik yang **terverifikasi**.
Bandingkan dengan protokol **yang sama**.
- **Split (train/valid/test)**
Pembagian data. Tulis **proporsi** dan apakah **official split**.

T

- **Throughput / Waktu (s)**
Kecepatan inferensi atau waktu proses.
Penting untuk aplikasi real-time.

U

- **Underfitting**
Model belum belajar pola.
Gejala: akurasi rendah di train dan test.
Solusi: model lebih kaya, fitur lebih kuat, training lebih lama.

V

- **Validation**
Set untuk *tuning*. **Jangan** pakai untuk *claim* akhir.
Gunakan *test set* untuk angka final.
- **Variance (SD)**
Ukuran sebaran hasil antar *fold/seed*.
Laporkan: mean $\pm$ **SD** atau **CI**.

W

- **Workflow “Collect → Dedupe → Tag → Cite → Format”**
Alur manajemen referensi di Zotero/Mendeley/EndNote.
Output: sitasi rapi & konsisten.

Referensi

Nature Editorial. (2019). *Reproducibility and data availability*. *Nature*, 569(7756), 5. <https://doi.org/10.1038/d41586-019-01307-2>

OpenAI. (2024). *GPT-4 Technical Report*. arXiv preprint arXiv:2303.08774. <https://arxiv.org/abs/2303.08774>

Purbo, O. W. (2025). *Menulis Jurnal di Bantu AI: Data-First Anti-Halusinasi*. Tangerang Selatan: Institut Teknologi Tangerang Selatan (ITTS) & OnnoCenter. Lisensi Creative Commons Attribution-ShareAlike 4.0 International. <https://creativecommons.org/licenses/by-sa/4.0/>

Raff, E. (2019). A step toward quantifying independently reproducible machine-learning research. *Communications of the ACM*, 63(4), 81–88. <https://doi.org/10.1145/3369750>

Stodden, V., Seiler, J., & Ma, Z. (2018). An empirical analysis of journal policy effectiveness for computational reproducibility. *Proceedings of the National Academy of Sciences*, 115(11), 2584–2589. <https://doi.org/10.1073/pnas.1708290115>

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., et al. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3(1), 160018. <https://doi.org/10.1038/sdata.2016.18>

Zotero. (2024). *Zotero reference management software [Computer software]*. Roy Rosenzweig Center for History and New Media, George Mason University. <https://www.zotero.org/>

Tentang Penulis



Onno W. Purbo, saat ini bertugas sebagai rektor di Institut Teknologi Tangerang Selatan (ITTS). Onno memperoleh gelar Ph.D bidang Electrical Engineering dari University of Waterloo, Canada, adalah seorang copyleft, educator dan ICT evangelist. Dia sudah mempublikasikan 50++ buku, termasuk free ICT ebook untuk sekolah tahun 2008. Beberapa buku terakhirnya adalah "Internet-TCP/IP: Konsep Dan Implementasi", 2018; "Sistem Operasi, Konsep Dan Membuat Linux OpenWRT Dan ROM Android", 2019; "IPv6 Untuk Mendukung Operasi Jaringan Dan Domain Name System", 2019; "Kubernetes untuk Pemula", 2024 dan "Membuat Operator Seluler 5G Sendiri", 2024. Dia memimpin sambungan pertama Internet di Institut Teknologi Bandung, tahun 1993-2000, dan menggunakannya untuk membuat jaringan Internet pendidikan yang pertama di

Indonesia. Dia membebaskan frekuensi WiFi, memperkenalkan RT/RW-net, antenna Wajanbolic dan jaringan selular OpenBTS dan private 5G sendiri. Dia memimpin jaringan telepon pertama di atas Internet, VoIP Merdeka, yang kemudian hari dikenal sebagai VoIP Rakyat berbasis SIP dan menggunakan kode area +62520 dan +62521. Dia saat ini aktif memperkenalkan e-Learning, dan menjalankan server e-Learning di <http://lms.onnocenter.or.id/moodle/> 66,000++ siswa / mahasiswa dan <https://opencourse.itts.ac.id> 39.000+ mahasiswa secara gratis. Email: onno@indo.net.id