

WORKFLOW

PENGEMBANGAN MODEL

Aditya Firman Ihsan

Misalkan pada suatu hari,
ibu menyuruhmu...

Ke supermarket, beli susu UHT 1 L

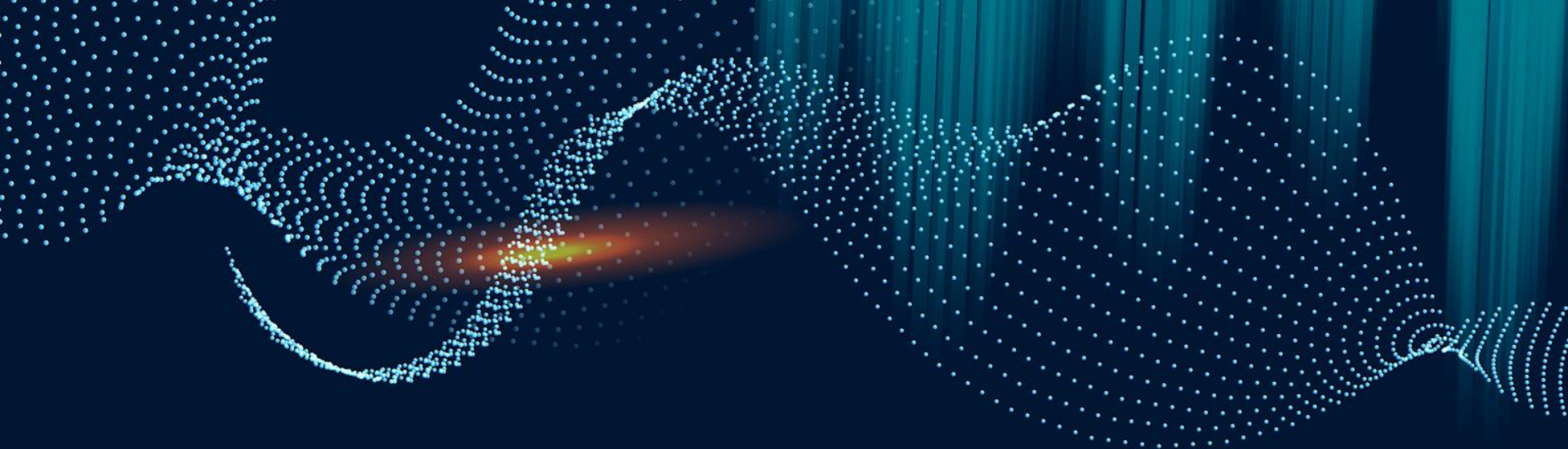
Misalkan pada suatu hari,
ibu menyuruhmu...

**Ke supermarket,
beli Heavenly Blush**

Misalkan pada suatu hari,
ibu menyuruhmu...

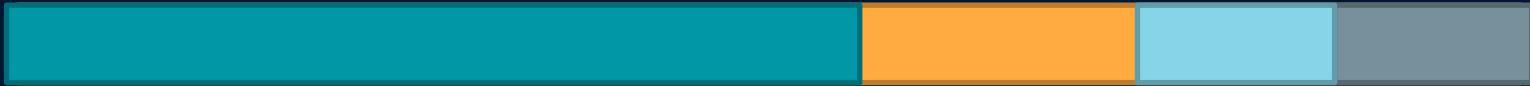
**Ke supermarket,
beli bahan untuk bikin pancake**

Bagaimana kita melakukan
sesuatu, bergantung seberapa
paham kita akan sesuatu itu



Menulis
program

Improve

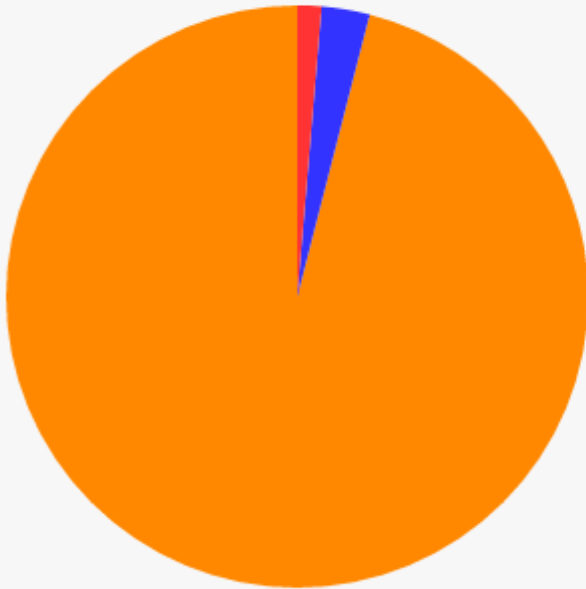


Memahami Masalah

Debugging

Programming help

- Ask other people
- Textbook
- Stack overflow



**Kita butuh paham
apa yang kita
lakukan untuk
bisa copy-paste
code yang tepat**



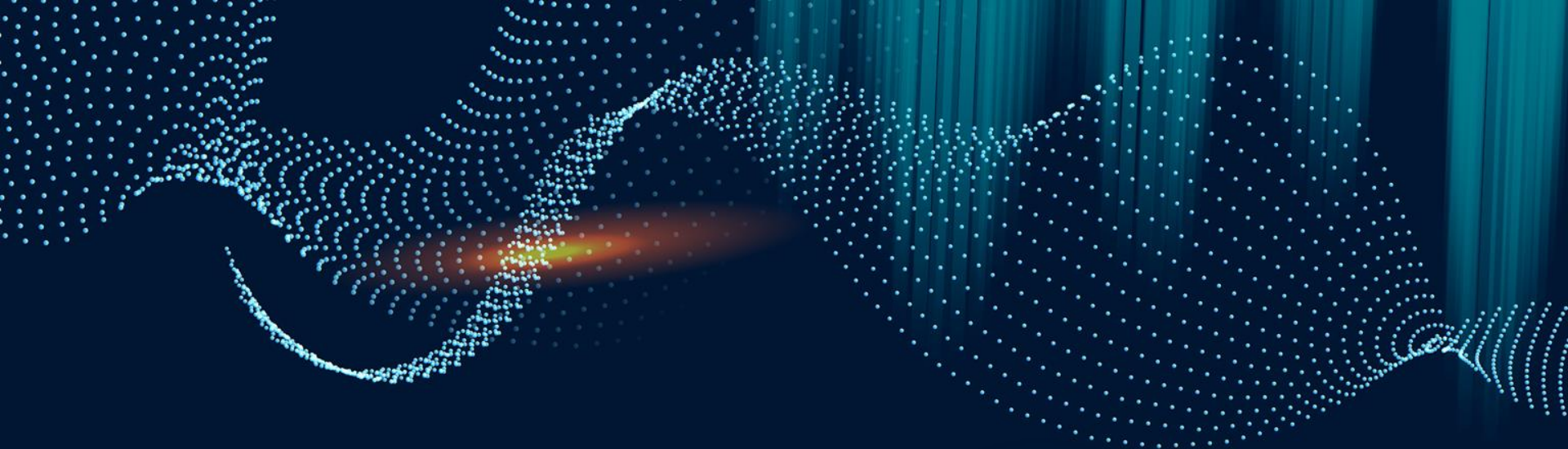
Jack Forge

@TheJackForge

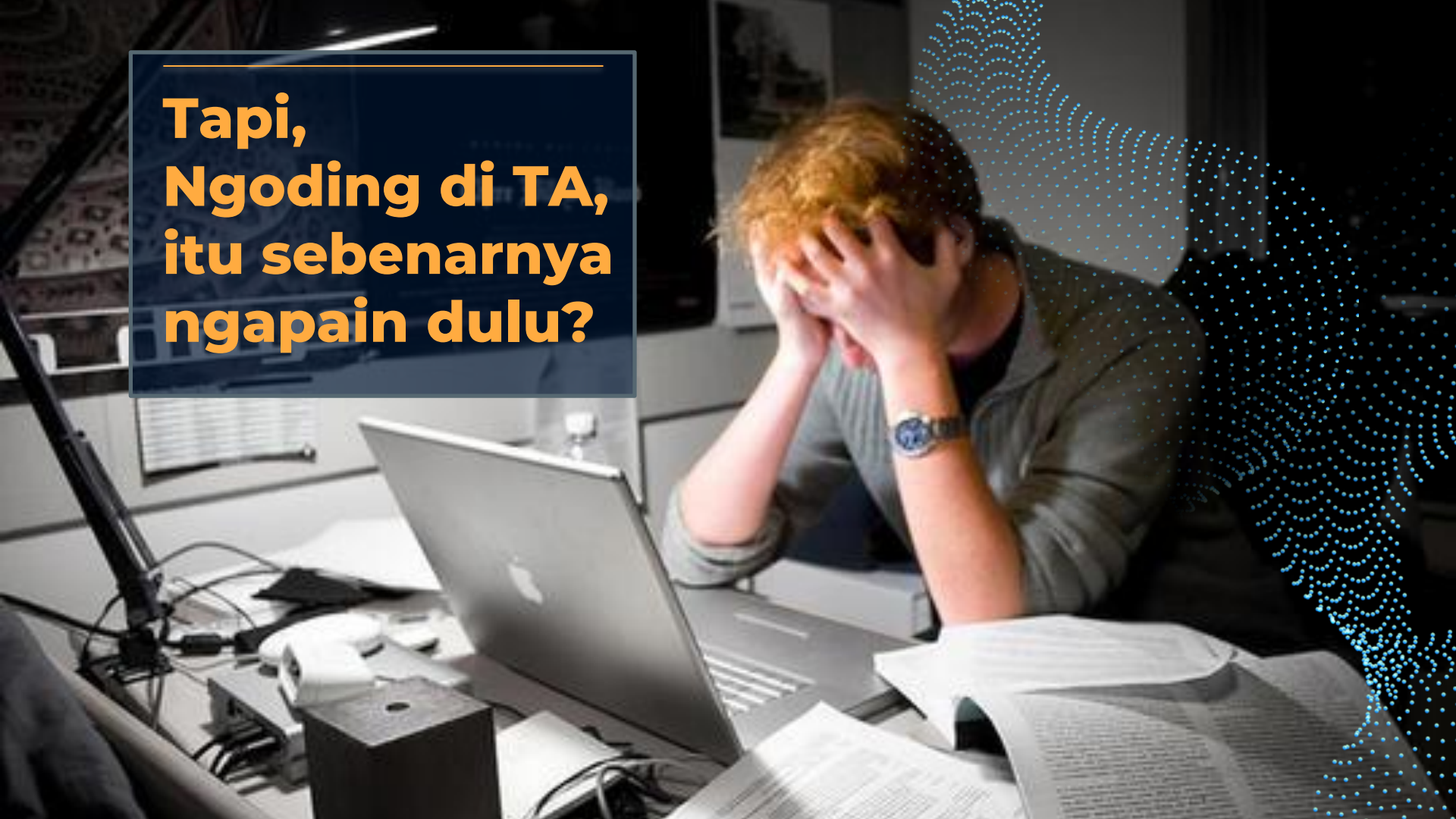
I'm a Pull Stack Developer.

I just pull things off the Internet
and put it into my code.

8:13 am · 19 Dec 20 · [Twitter for Android](#)



**Tapi,
Ngoding di TA,
itu sebenarnya
ngapain dulu?**



**Apa sebenarnya
yang dilakukan
ketika riset?**



Penelitian

- 1. n pemeriksaan yang teliti; penyelidikan**
- 2. n kegiatan pengumpulan, pengolahan, analisis, dan penyajian data yang dilakukan secara sistematis dan objektif untuk memecahkan suatu persoalan atau menguji suatu hipotesis untuk mengembangkan prinsip-prinsip umum**

Penelitian

1. n pemeriksaan yang teliti;



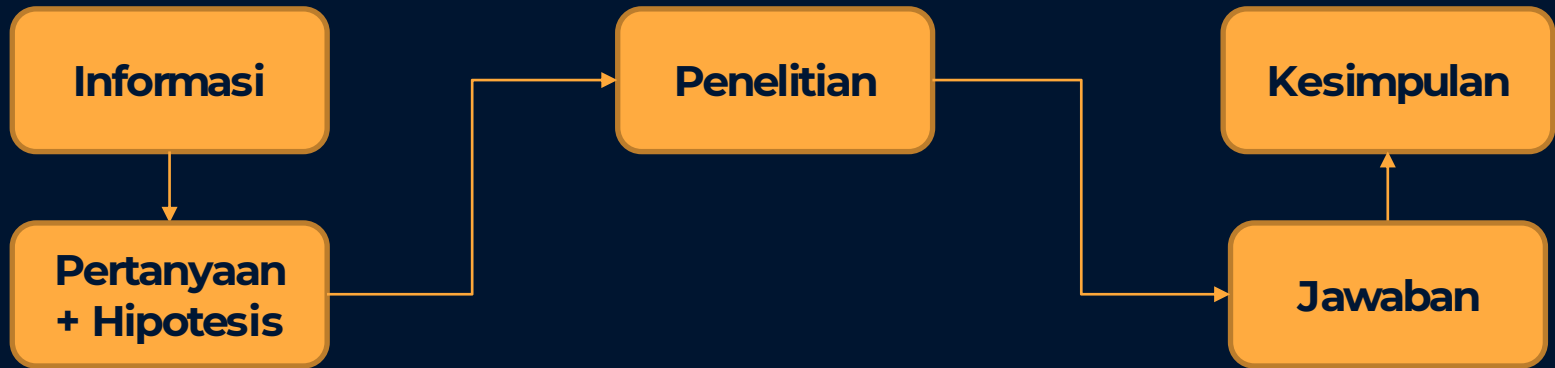
prinsip umum

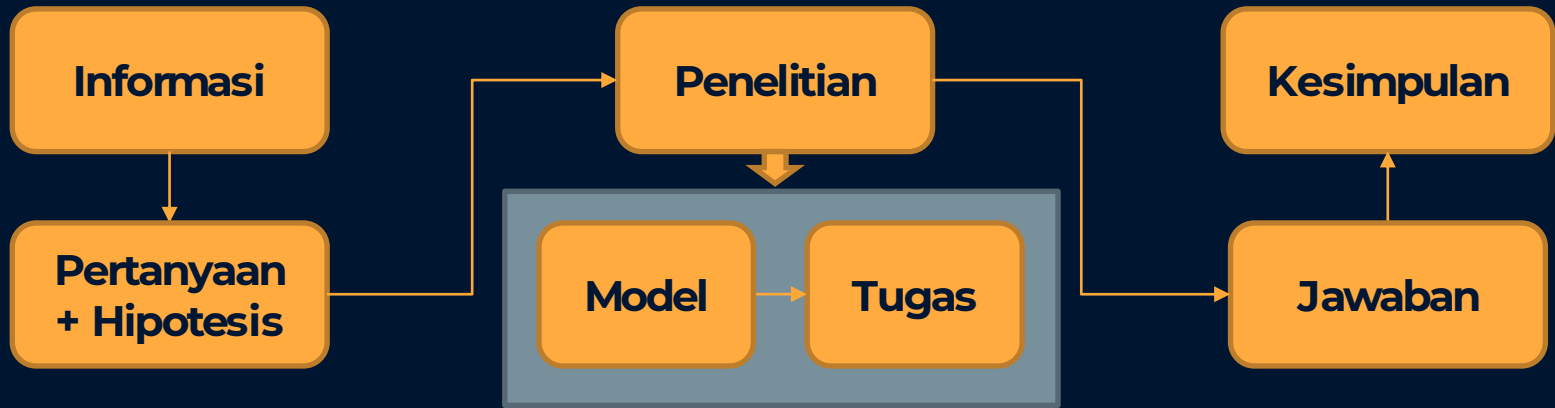
```
graph LR; A[Informasi] --> B[Penelitian]; B --> C[Kesimpulan]
```

Informasi

Penelitian

Kesimpulan





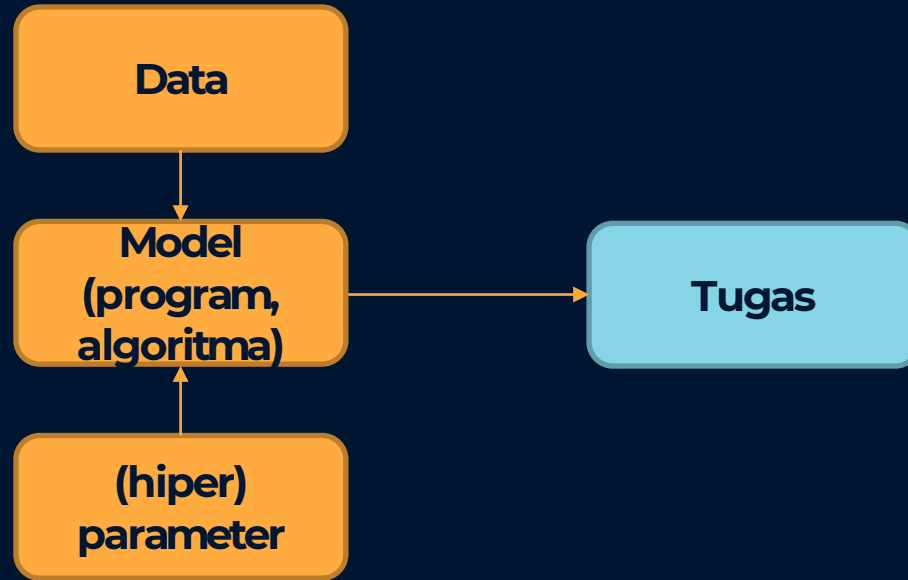
**Ada apa dengan
model?**

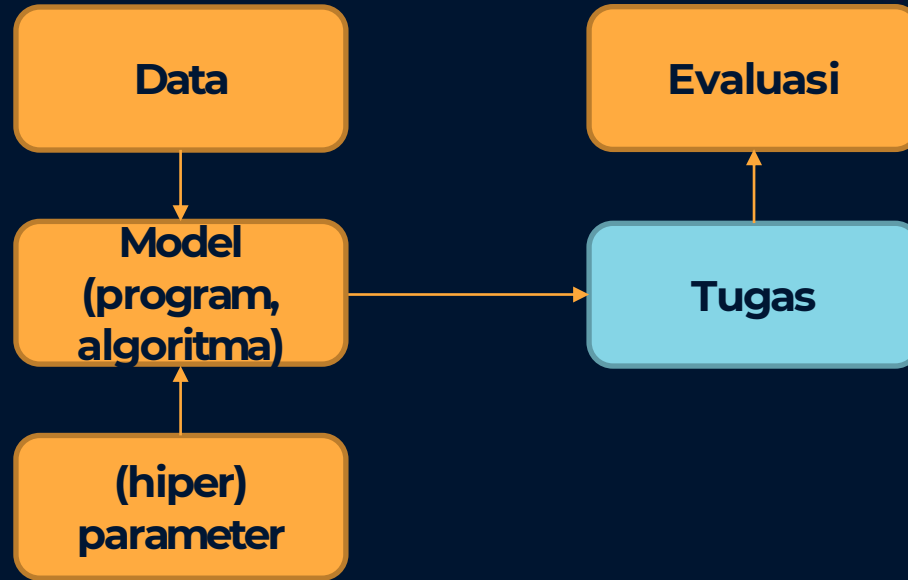


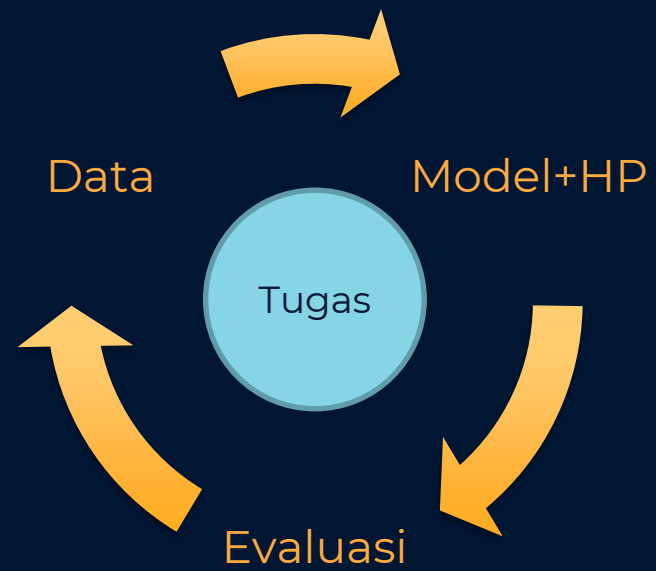
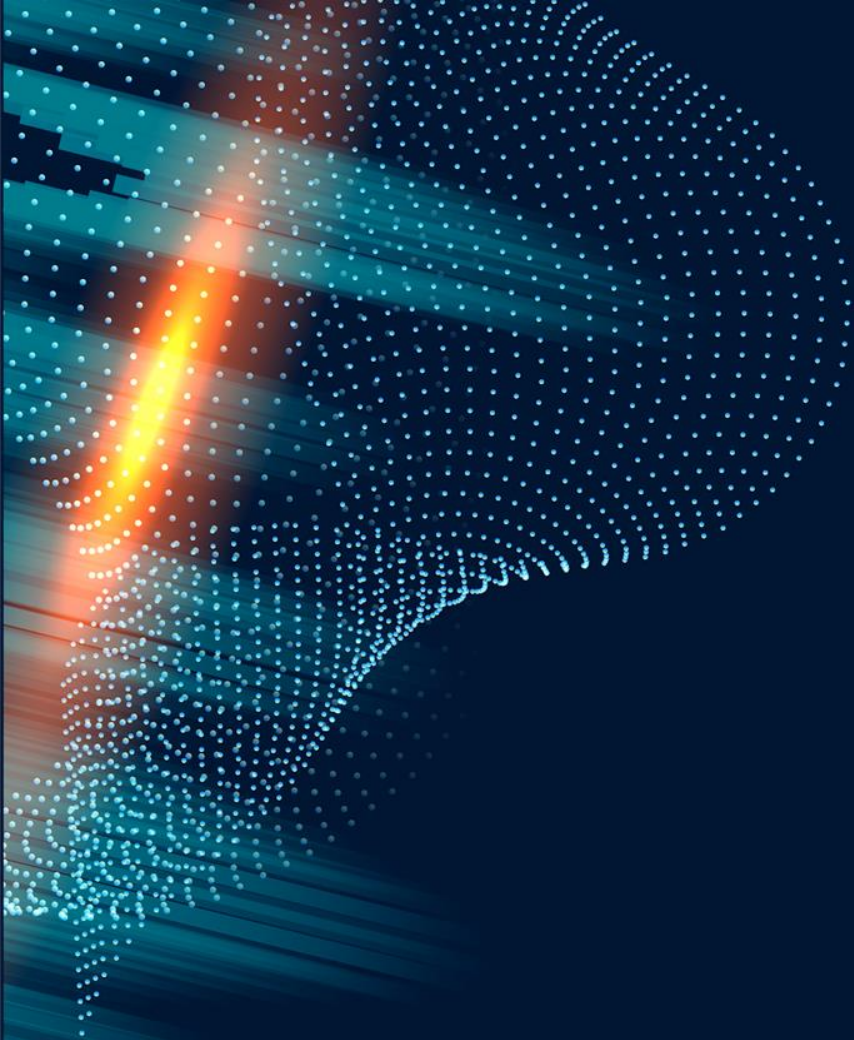
**Model
(program,
algoritma)**

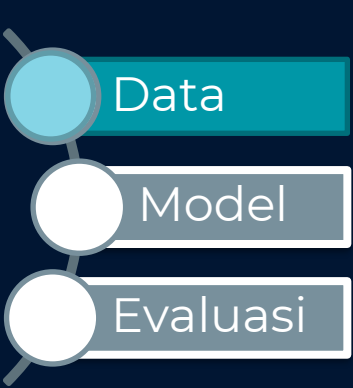


Tugas





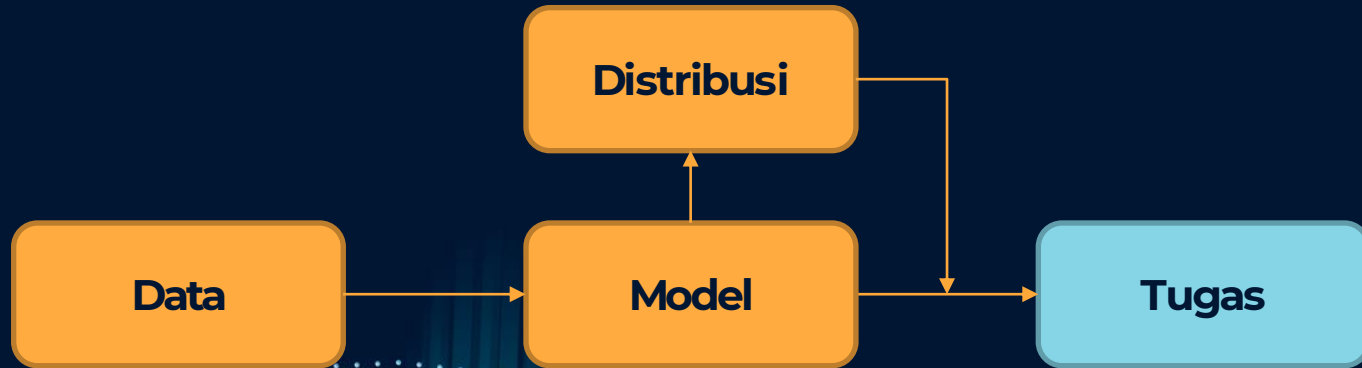




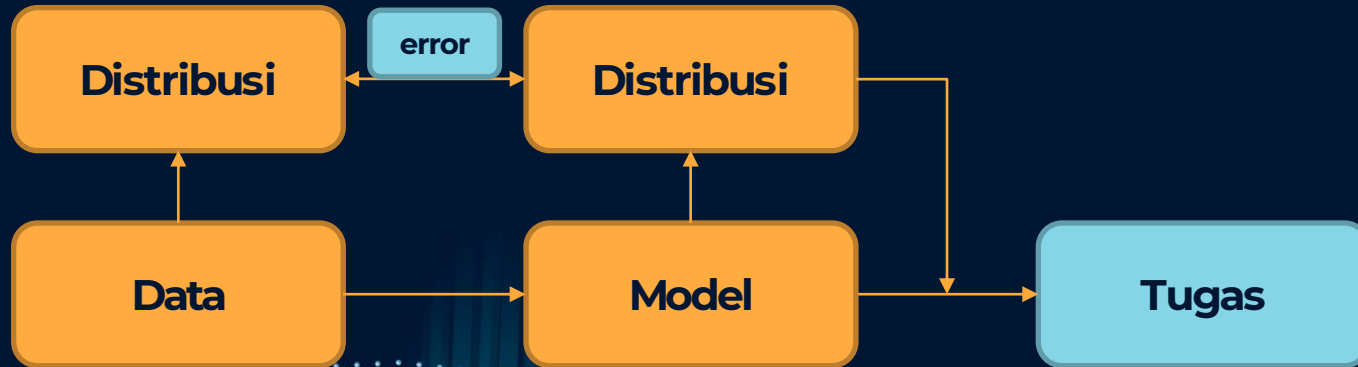
Mengatur Data



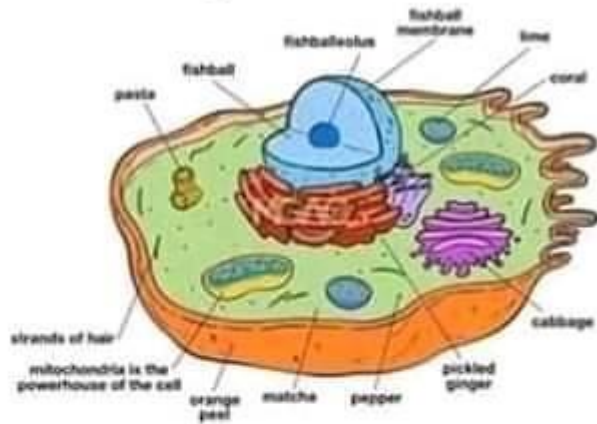
Yang sebenarnya dilakukan *learning model* adalah membaca distribusi dari suatu data



Distribusi yang dibaca model hanya aproksimasi



What we study in class:

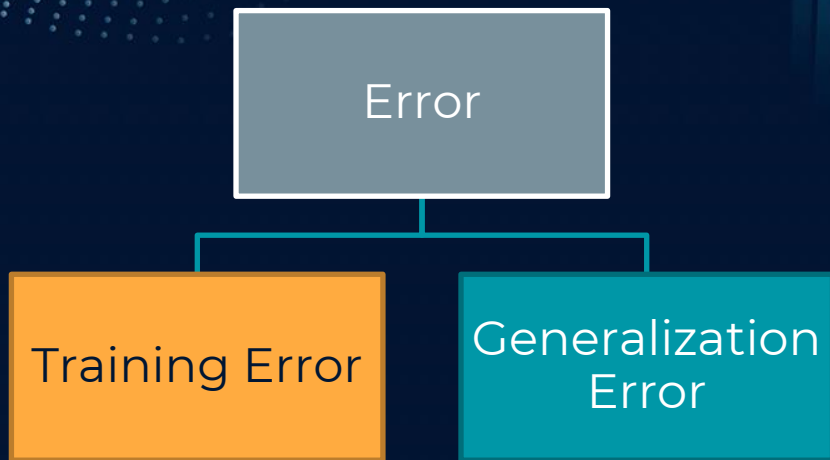


What we see on the exam paper:



Sayangnya,

Distribusi suatu data dari yang dipelajari dengan yang ditemui ketika diaplikasikan bisa berbeda

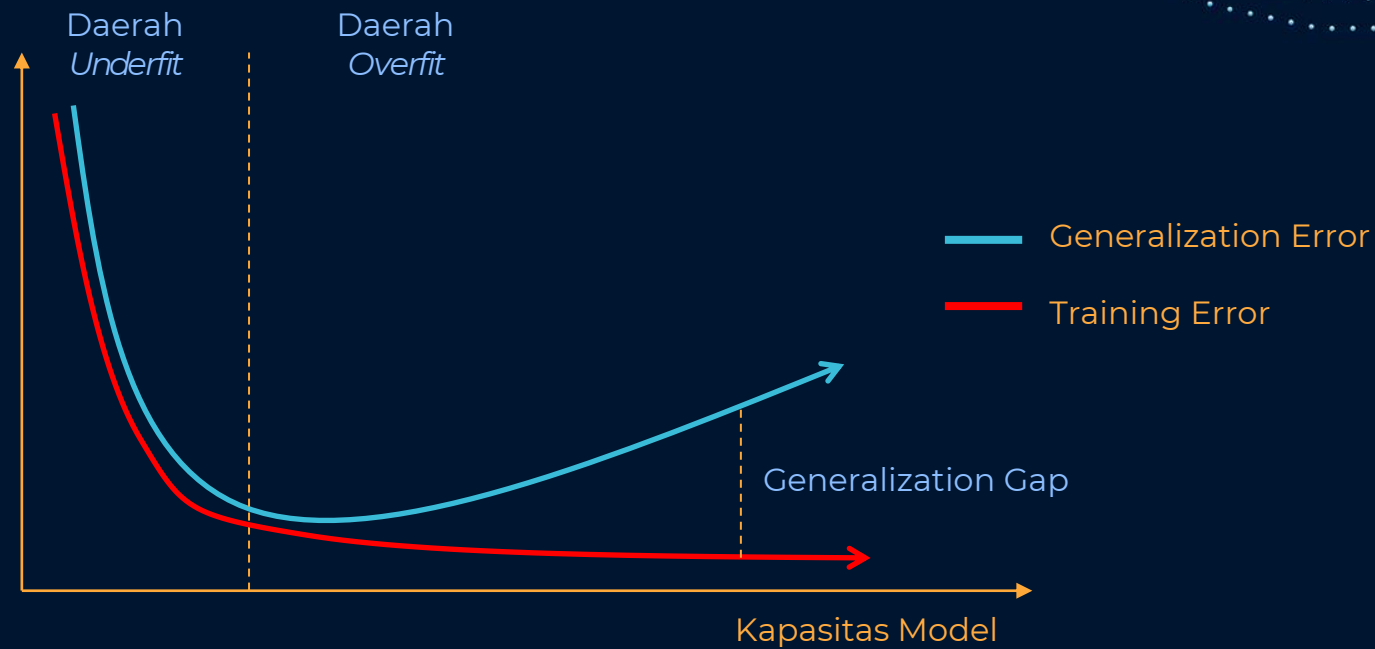


Training set

Validation/Development set



Generalization Gap



Berolahraga dan minum vitamin, agar tetap fit

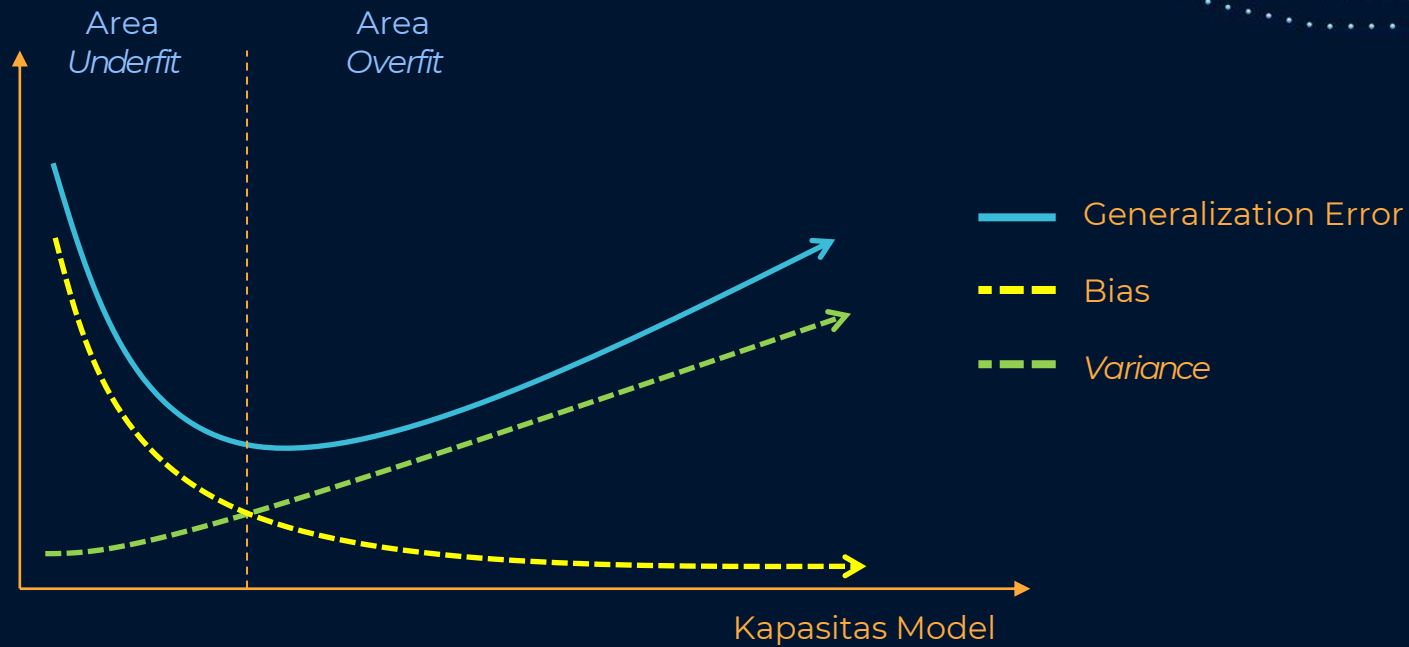
Underfit

- Model terlalu sederhana, sehingga terlalu mudah mengambil keputusan
- Terbawa prasangka, kurang variatif
- Bias tinggi, variansi rendah

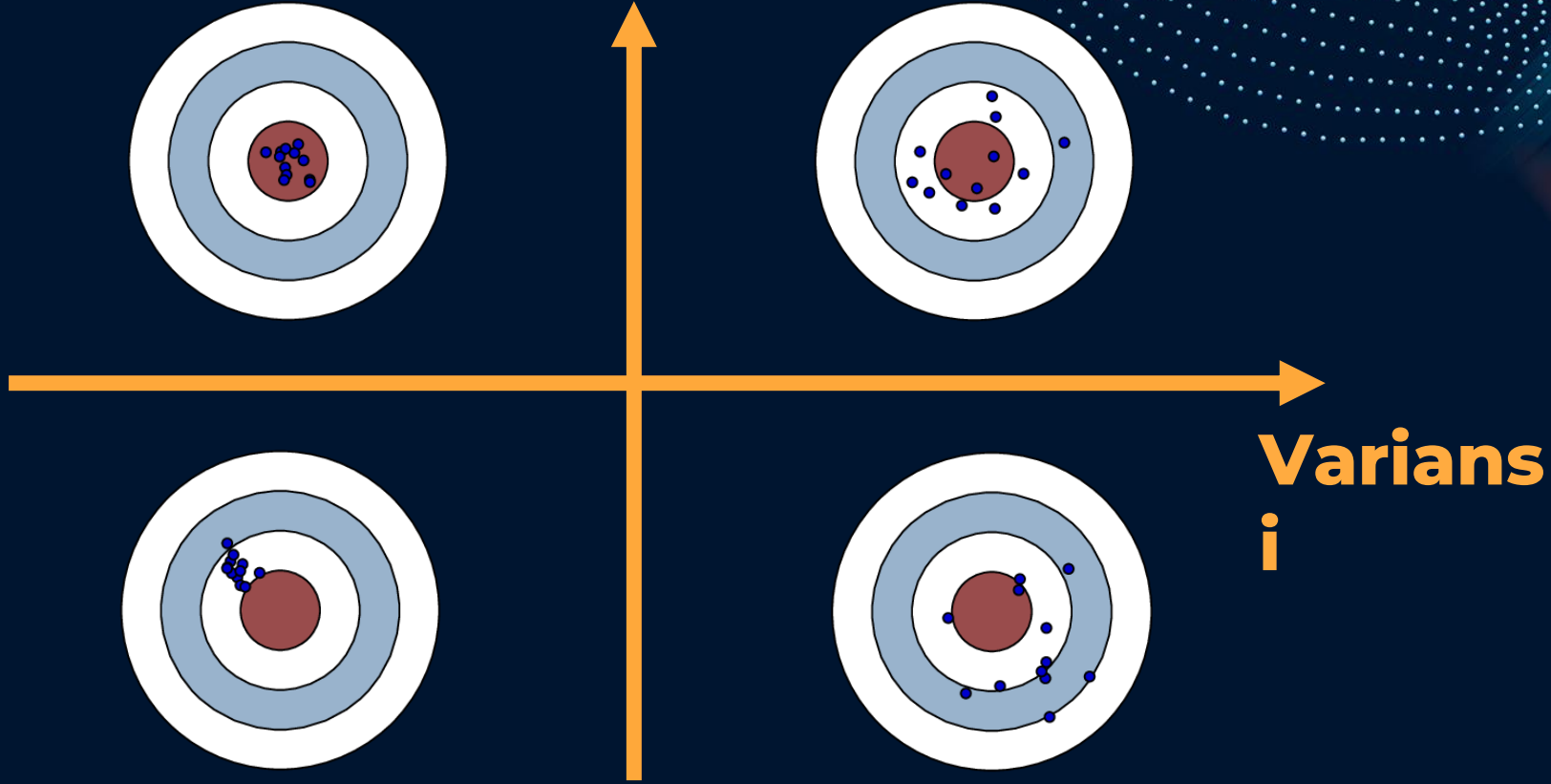
Overfit

- Model terlalu kompleks, terlalu fokus pada hal detail
- Terpengaruh data, tidak punya pendirian
- Bias rendah, variansi tinggi

Bias dan Variansi



Bias



Varians
i



Terus harus ngapain?

Bias tinggi

- Butuh model lebih kompleks
- Latihan lebih lama

Variansi Tinggi

- Butuh model lebih sederhana
- Latihan lebih singkat
- Regularisasi
- Lebih banyak data

Data harus dipastikan cukup, terutama validation set untuk mengukur variansi

Training set

Validation/Development set

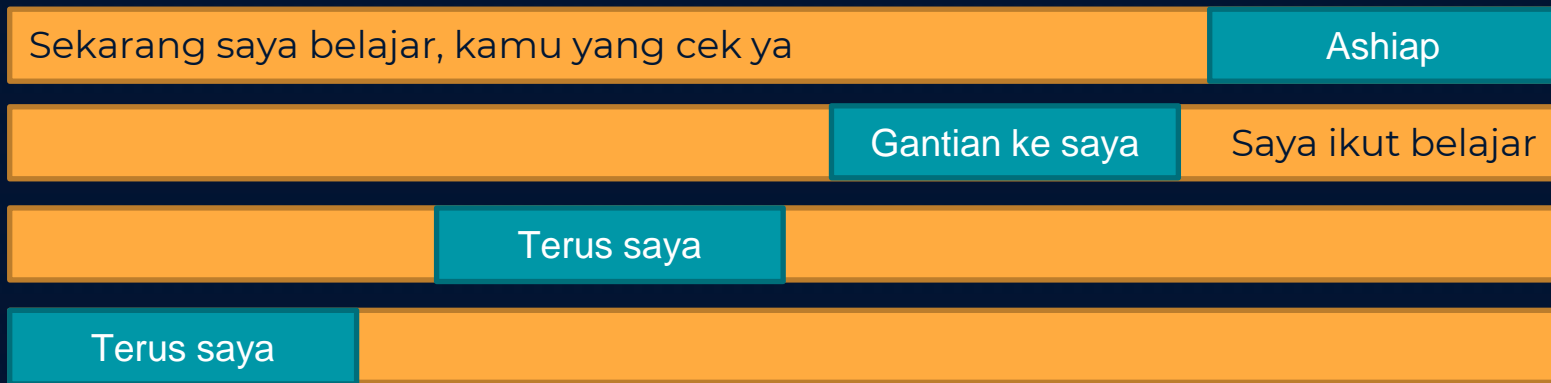


Kalau sedikit bagaimana?

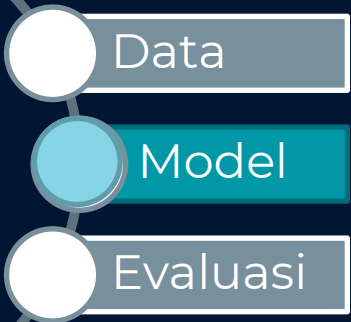
Saling kroscek saja antar data!

Training set

Validation/Development set



Nama kroscek ga keren, jadilah **cross-validation**



Mempercantik Model





Yang butuh di-tuning

**Bagaimana cara mencari
jodoh parameter yang baik?**





01

**Tentukan metrik-
metrik utama**

02

**Jika lebih dari 1,
maka gabungkan**

03

**Tetapkan batas cukup
pada metrik lainnya**



1


**Daftarkan dulu
hiperparameter yang
dimiliki model**

2

**Pilih sebagian yang
kiranya sesuai**

The background features a dark blue gradient with a series of parallel, slightly curved lines in shades of teal and blue. A bright, glowing orange and yellow light trail curves across the upper left portion of the image. A dense field of small, light blue dots is scattered across the scene, particularly concentrated in the upper right and lower left areas, creating a sense of depth and complexity.

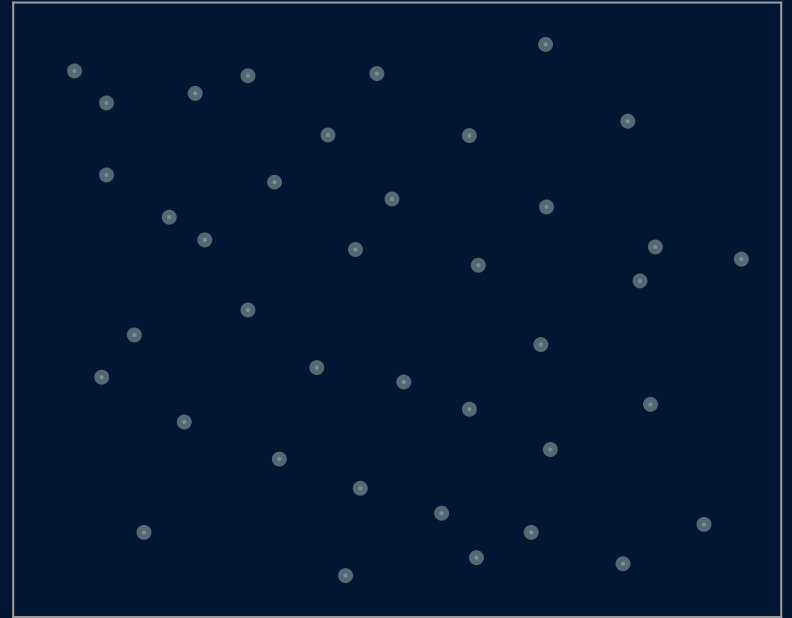
**Tapi, bagaimana
kalau hyperparameter
yang dicari banyak?**

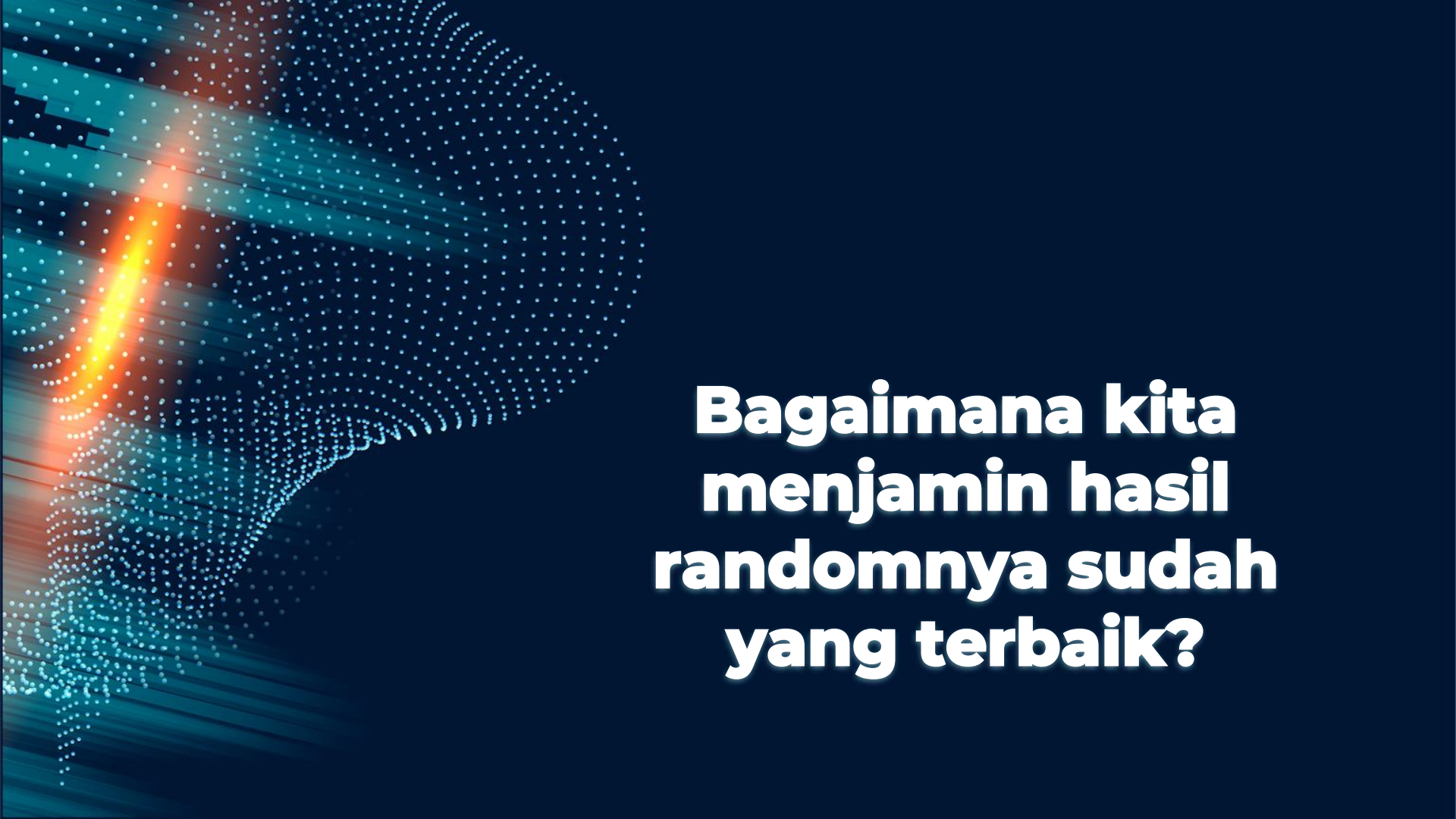


Ya diseleksi saja.
Pilih 2 yang paling
penting

**Sayangnya, kita tidak
tahu parameter yang
mana**

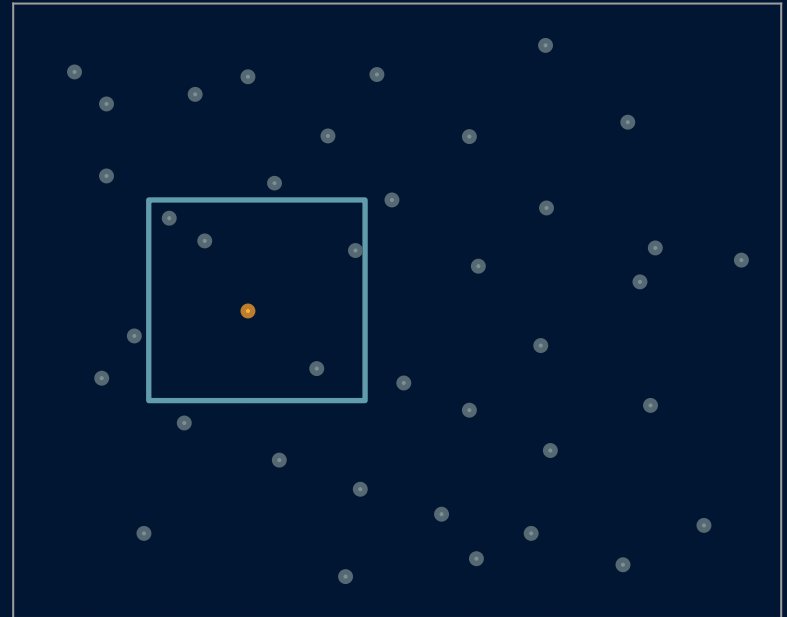
Random Search



The background features a dark blue gradient with a grid of small, glowing blue dots that form a curved, funnel-like shape. A bright orange and yellow light trail streaks across the left side, creating a lens flare effect.

**Bagaimana kita
menjamin hasil
randomnya sudah
yang terbaik?**

Perbaiki domain



3

Bila hiperparameternya sedikit, pakai Grid Search

4

Bila banyak, pakai Random Search secara bertahap

*

Identifikasi skala setiap hiperparameter

learning rate

0.01 ————— 0.05

10^{-4} ————— 10^{-1}

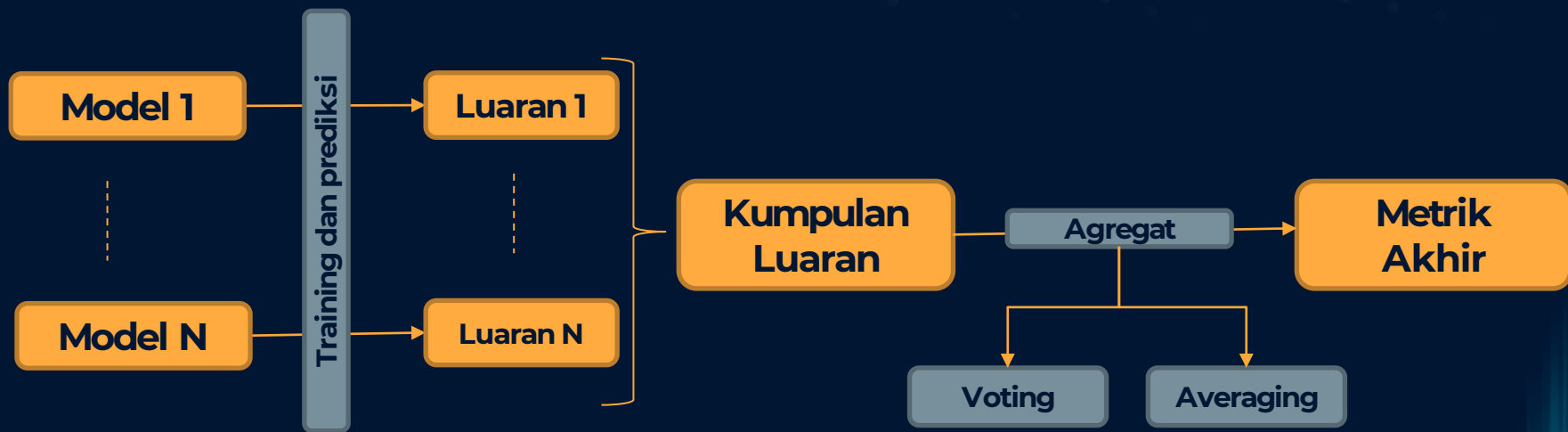
Hiperparameter cukup?



Avengeeers! Ensemble!



Ensemble: Belajar bersama itu lebih baik



Pembagian model bisa berdasarkan: **data, fitur, hyperparameter, arsitektur**

**Belajar lah dari
yang sudah
berpengalaman**



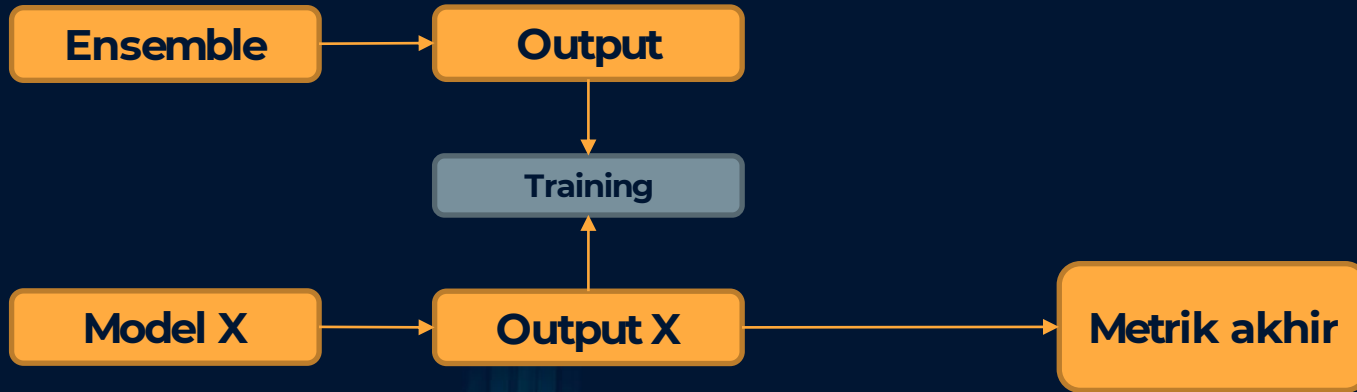
**KNOWLEDGE
DISTILLATION**



**SELF-
DISTILLATION**

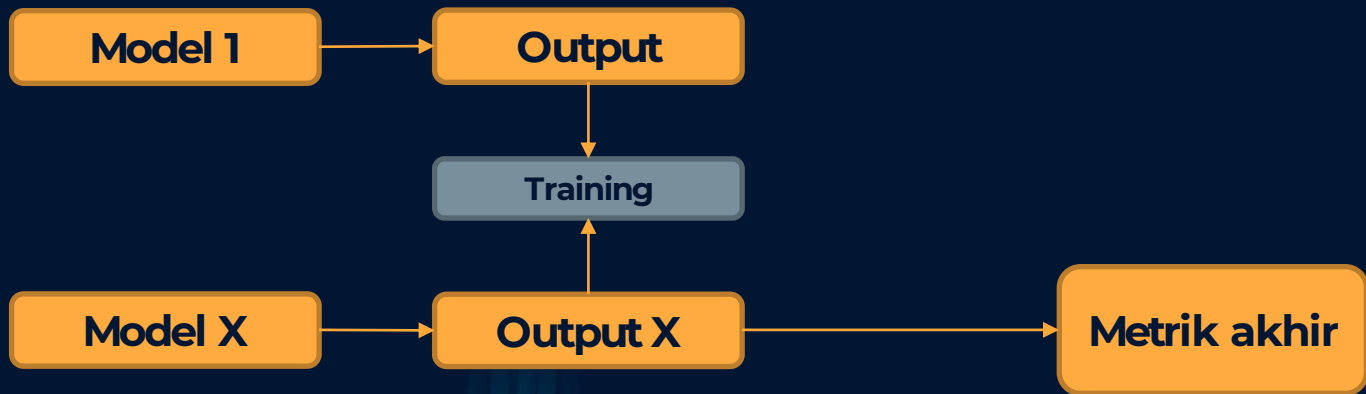
Knowledge Distillation

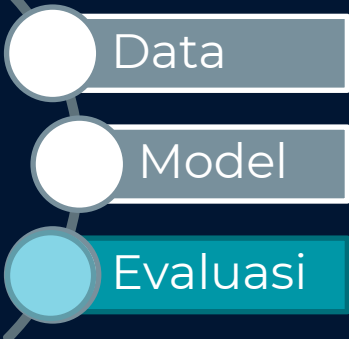
Memiliki model gabungan bisa berat ketika diaplikasikan



Self Distillation

Bagaimana kalau cukup belajar dari satu model?

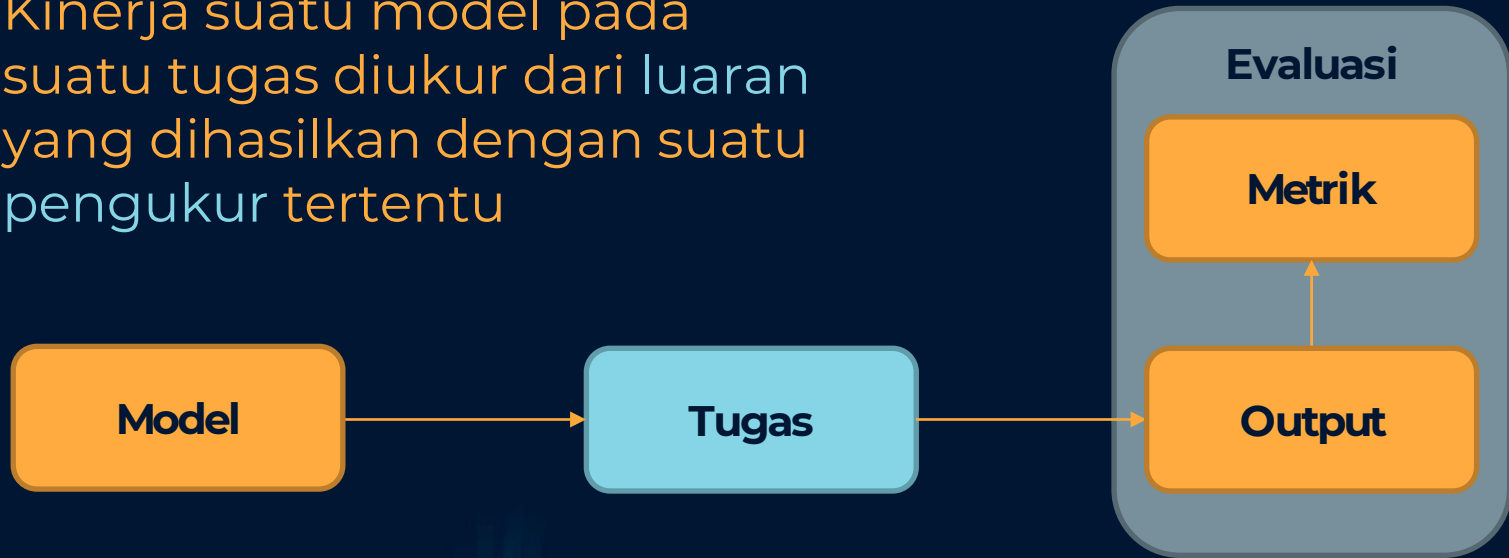




Mengukur keberhasilan



Kinerja suatu model pada suatu tugas diukur dari luaran yang dihasilkan dengan suatu pengukur tertentu



Evaluasi


```
graph TD; A[Evaluasi] --> B[Tentukan Metrik yang tepat]; A --> C[Optimalkan metriknya]; B --- D[Goal Setting]; C --- E[Error Analysis];
```

Tentukan
Metrik yang
tepat

Goal Setting

Optimalkan
metriknya

Error Analysis



**Selalu gunakan
“single number
evaluation metric”**

Untuk menciptakan 1 goal
yang menjadi fokus

Misal kita ingin mengukur seberapa bagus kinerja perguruan tinggi

	Impact Rank	Openness Rank	Excellence Rank
ITB	749	6492	1392
IPB	592	987	2175
Tel-U	692	1773	2653
UGM	594	775	1503
ITS	784	1427	1741

Misal kita ingin mengukur seberapa bagus kinerja perguruan tinggi

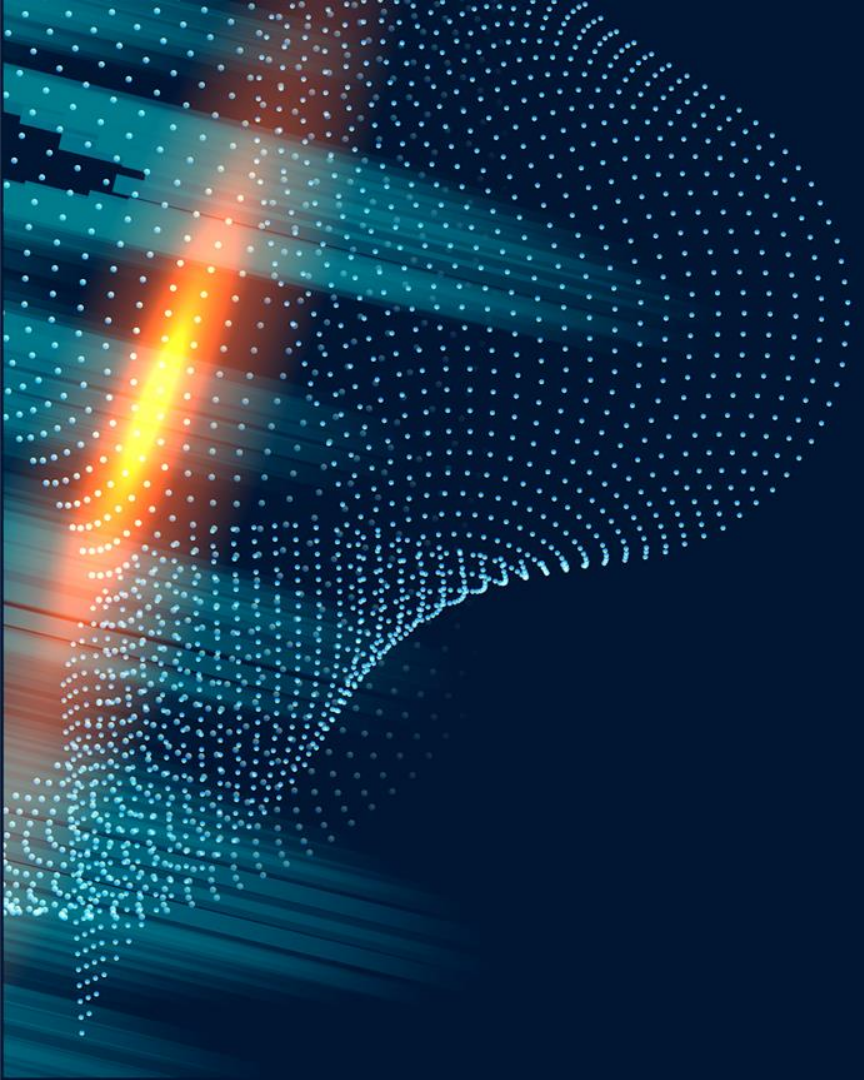
	Impact Rank	Openness Rank	Excellence Rank	Overall Rank
ITB	749	6492	1392	2126
IPB	592	987	2175	1092
Tel-U	692	1773	2653	1417
UGM	594	775	1503	852
ITS	784	1427	1741	1103

Ada perhitungan bobotnya

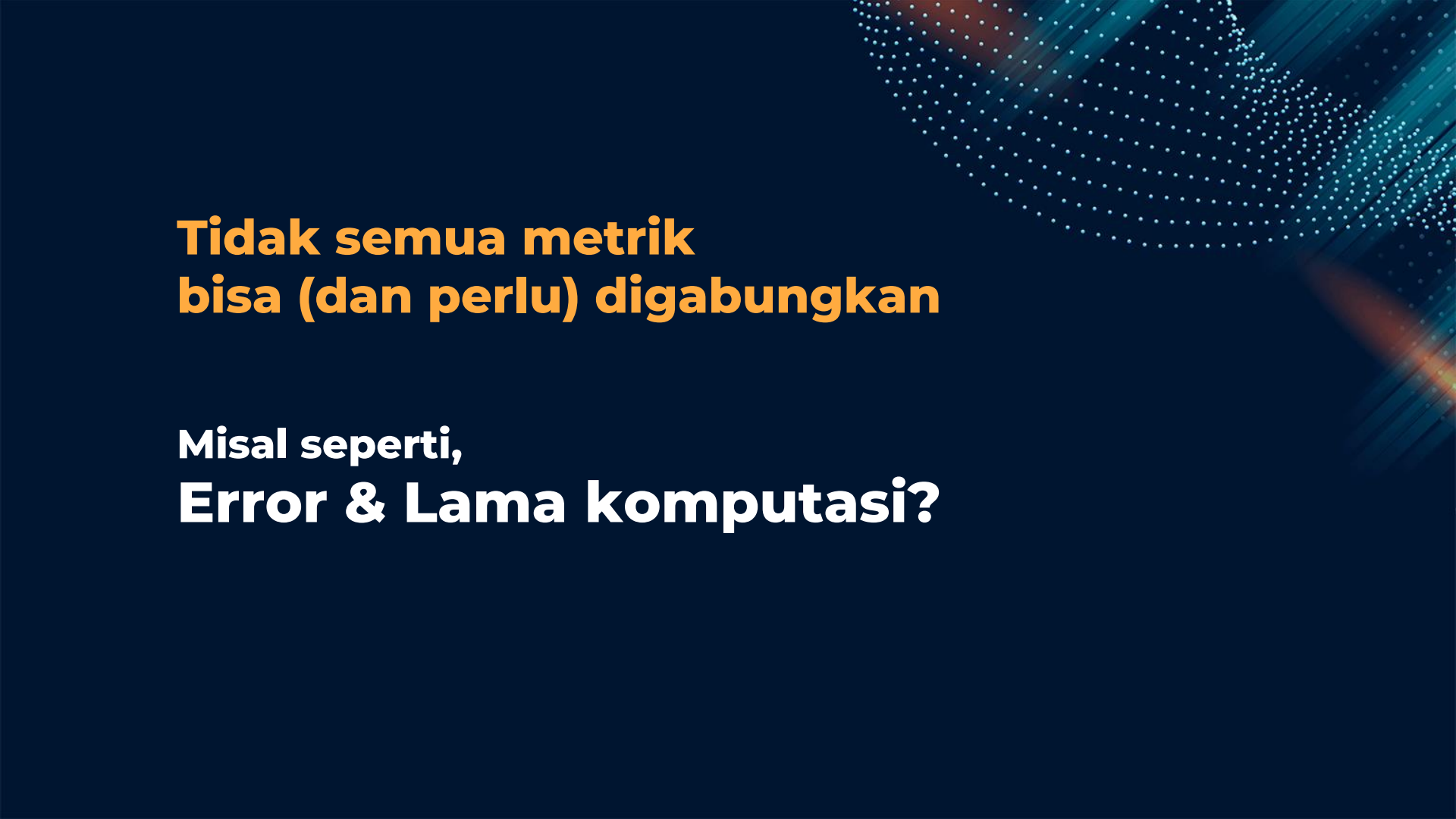
Metrik	Bobot
IMPACT	0.5
OPENNESS	0.1
EXCELLENCE	0.4

Demikian juga model

	Precision	Recall	Accuracy
Model A	0.97	0.92	0.94
Model B	0.90	0.91	0.98
Model C	0.93	0.96	0.91

The background features a dark blue gradient. On the left side, there are several horizontal, slightly curved light trails in shades of cyan and blue, with a bright orange and yellow glow at their ends. To the right of these trails, a dense field of small, glowing blue dots is arranged in a pattern that resembles a particle detector or a data visualization of a complex system.

**Kalau terdapat
banyak metrik
yang tidak
terkait
bagaimana?**



**Tidak semua metrik
bisa (dan perlu) digabungkan**

**Misal seperti,
Error & Lama komputasi?**

Metrik

```
graph TD; A[Metrik] --> B["Pengoptimal  
(Optimizing)"]; A --> C["Pencukup  
(Satisficing)"]; B --> D["Metrik yang jadi  
ukuran utama"]; D --> E["Dimaksimalkan/  
Diminimalkan"]; C --> F["Metrik pembatas  
sebagai ukuran cukup"]; F --> G["Dijaga dalam  
suatu batas"];
```

Pengoptimal
(*Optimizing*)

**Metrik yang jadi
ukuran utama**



**Dimaksimalkan/
Diminimalkan**

Pencukup
(*Satisficing*)

**Metrik pembatas
sebagai ukuran cukup**



**Dijaga dalam
suatu batas**

Nilai Error & Lama komputasi?

Opsi 1

Metrik Pengoptimal: Nilai Error

Metrik Pencukup: Lama Komputasi

Target: Perkecil Error seminimal mungkin dengan lama komputasi cukup < 1 menit

Opsi 2

Metrik Pengoptimal: Lama Komputasi

Metrik Pencukup: Nilai Error

Target: Perkecil waktu komputasi sesingkat mungkin dengan nilai error cukup < 0.1

Gagal menetapkan metrik yang tepat

Seperti berusaha mengukur potensi seorang anak dari nilai UN-nya,

Seperti berusaha menembak rusa tapi yang dibidik pohon di belakangnya

Seperti berusaha merebut hati seseorang tapi yang dideketi sahabatnya (duh)



YOU MISSED



THE TARGET



01

**Tentukan metrik-
metrik utama**

02

**Jika lebih dari 1,
maka gabungkan**

03

**Tetapkan batas cukup
pada metrik lainnya**

The background features a dark blue gradient with a series of parallel, slightly curved lines in shades of teal and blue. A bright orange and yellow light trail, resembling a comet or a laser beam, curves across the left side. A grid of small, glowing blue dots is scattered across the scene, particularly concentrated in the upper left and middle sections.

**Setelah
tentukan
metriknya,
selanjutnya?**

Sebelum lanjut,

Ini anjing atau kucing?



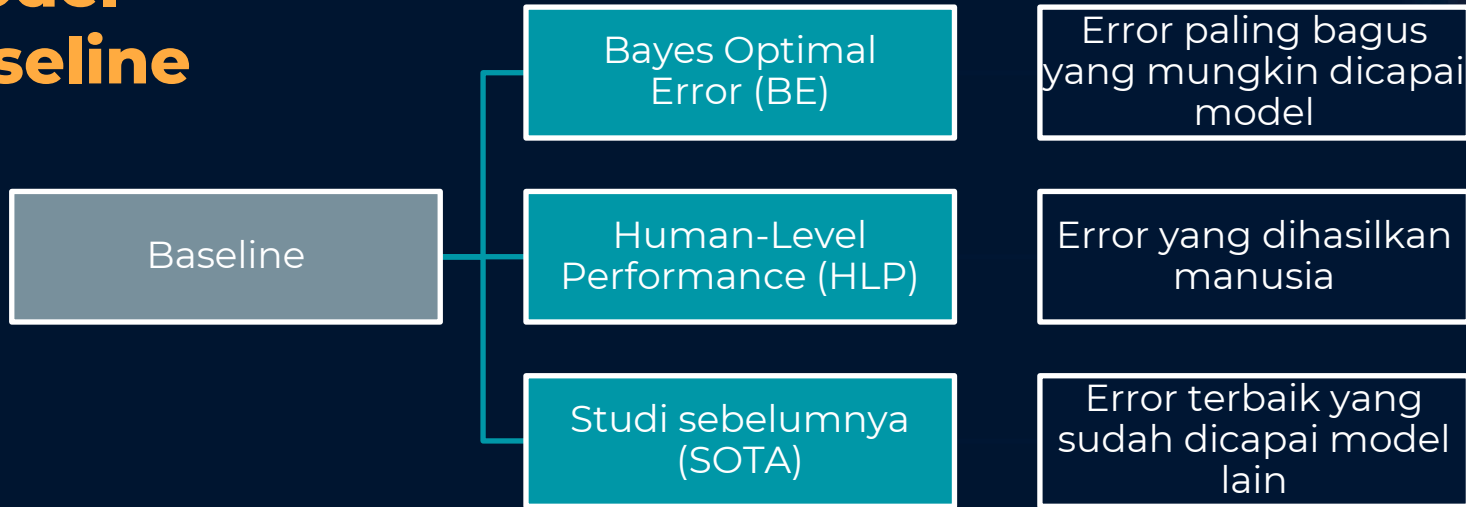


**Manusia diberi masalah
sesuai dengan batas
kemampuannya,**

Demikian juga model

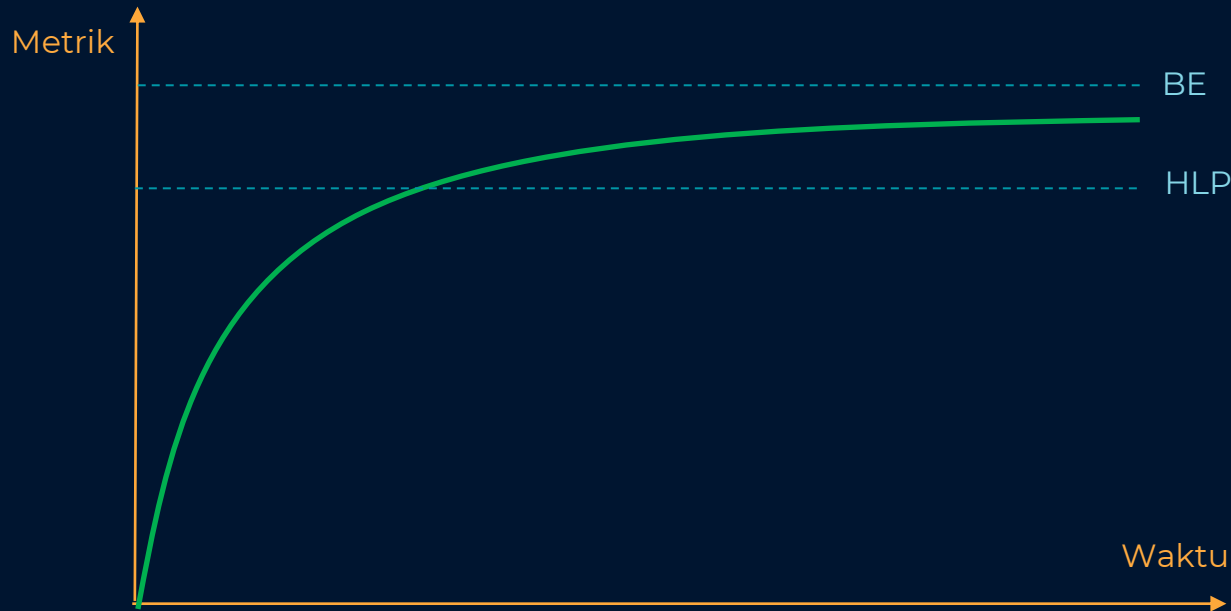
Kita harus tahu batas
paling baik suatu model

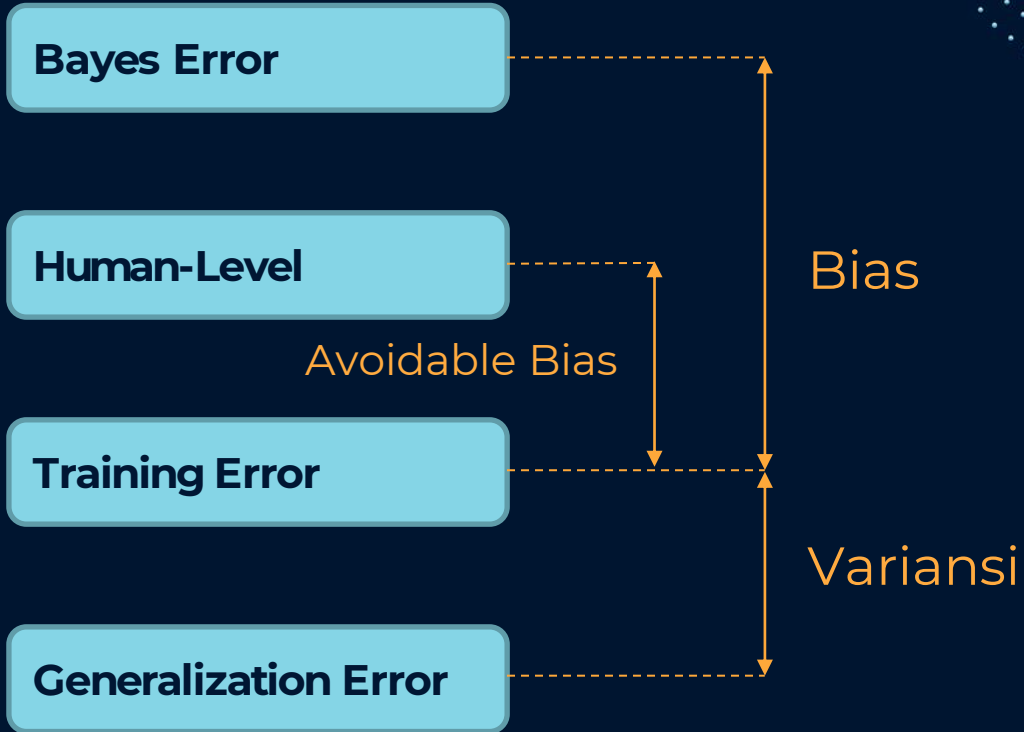
Patokan pengembangan model = baseline



$$\mathbf{BE \leq SOTA \leq HLP}$$

Kita pada dasarnya tidak bisa tahu BE, tapi bisa gunakan HLP sebagai pendekatan





Hate Speech atau Offensive?

“GA PERNAH MENDALAMI AL-QURAN
YA BANG??? PANTESAN MULUTNYA
KAYA ORANG KAFIR BEJAT HINA”

“Ya Bani Taplak dkk”

Setidaknya gw punya jari tenga buat lu,
sebelum gw ukur nyali ama bacot lu

Error Analysis

Periksa data-data yang membuat model gagal melaksanakan tugas

	Ciri 1	Ciri 2	Label keliru	Komentar
Sampel 1					
...					
Sampel n					
% total					

Error Analysis

Contoh data gambar

	Blur	Objek tidak di tengah	Kontras rendah	Label keliru	Komentar
Sampel 1		√			
...					
Sampel n				√	
% total	10%	21%	32%	1%	

Error Analysis

Contoh data suara

	Suara gemerisik	Bicara terlalu cepat	Suara naik turun	Label keliru	Komentar
Sampel 1		√			
...					Ada mobil lewat
Sampel n				√	
% total	2%	1%	5%	0.5%	

Error Analysis

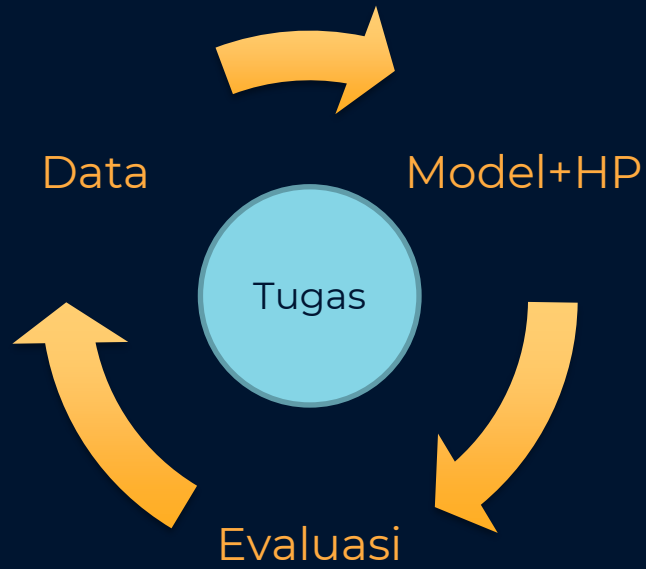
Contoh data teks

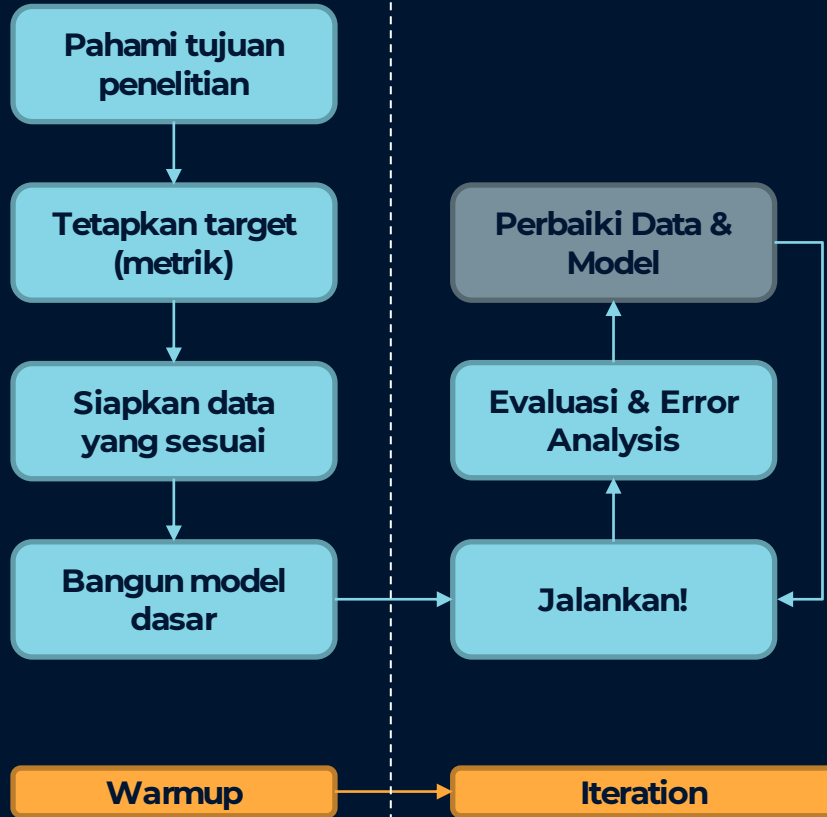
	Bahasa alay	Kata-kata ambigu	Kalimat <3 kata	Label keliru	Komentar
Sampel 1		√			
...					Isinya emoticon semua
Sampel n				√	
% total	9%	3%	1%	21%	

Fyuh...

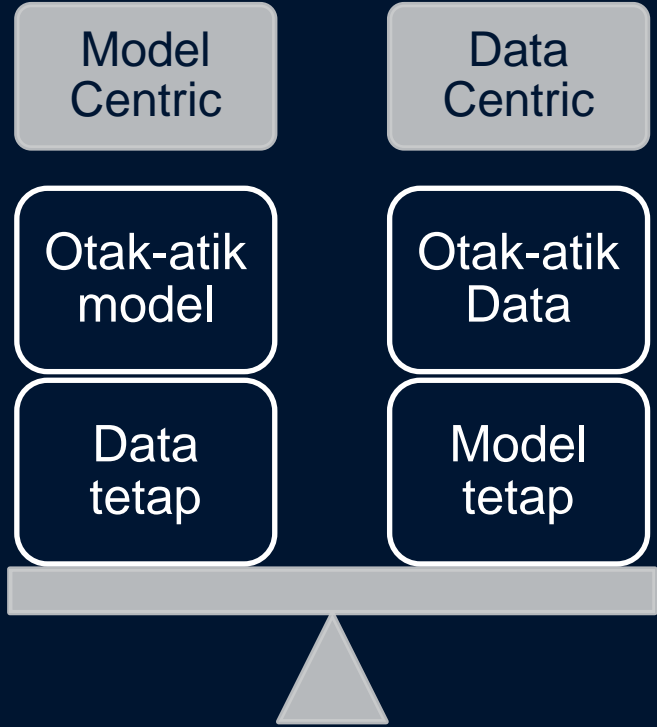
**Terus,
Intinya
bagaimana
pak?**

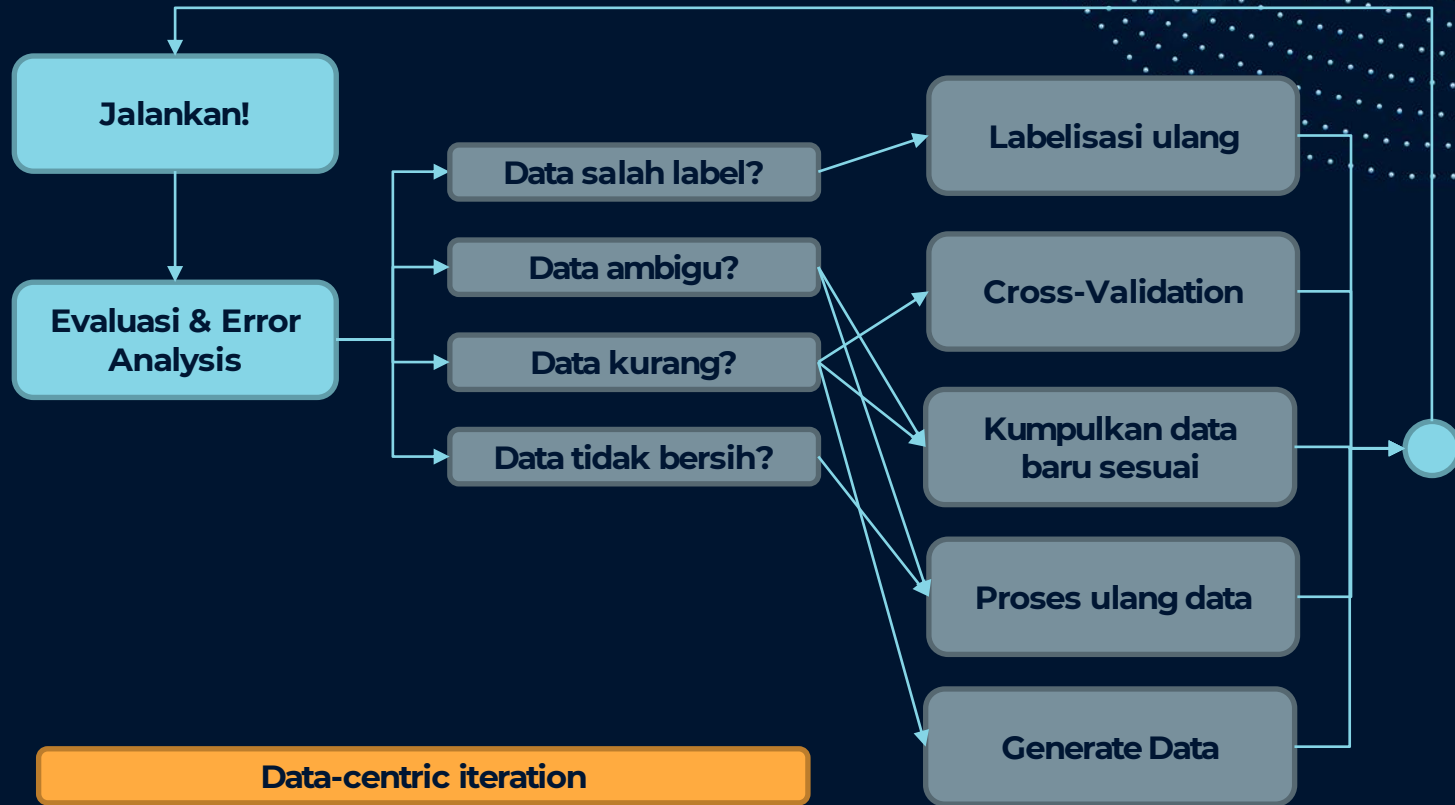


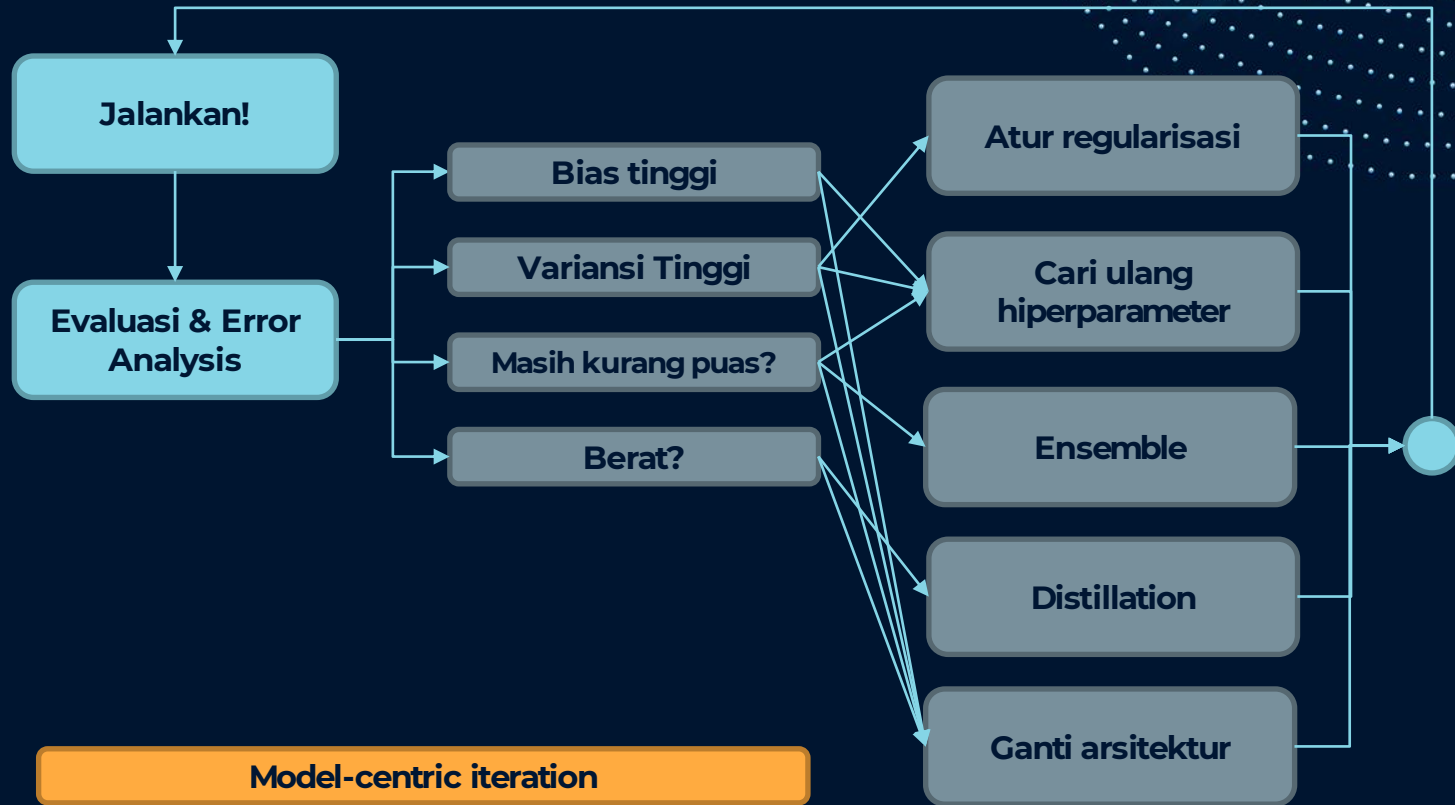




Kalau warm-up tidak dilakukan dengan baik, maka pasti keram waktu iterasi







Kembangkan perlahan, dari sederhana ke kompleks

NOT LIKE THIS



LIKE THIS



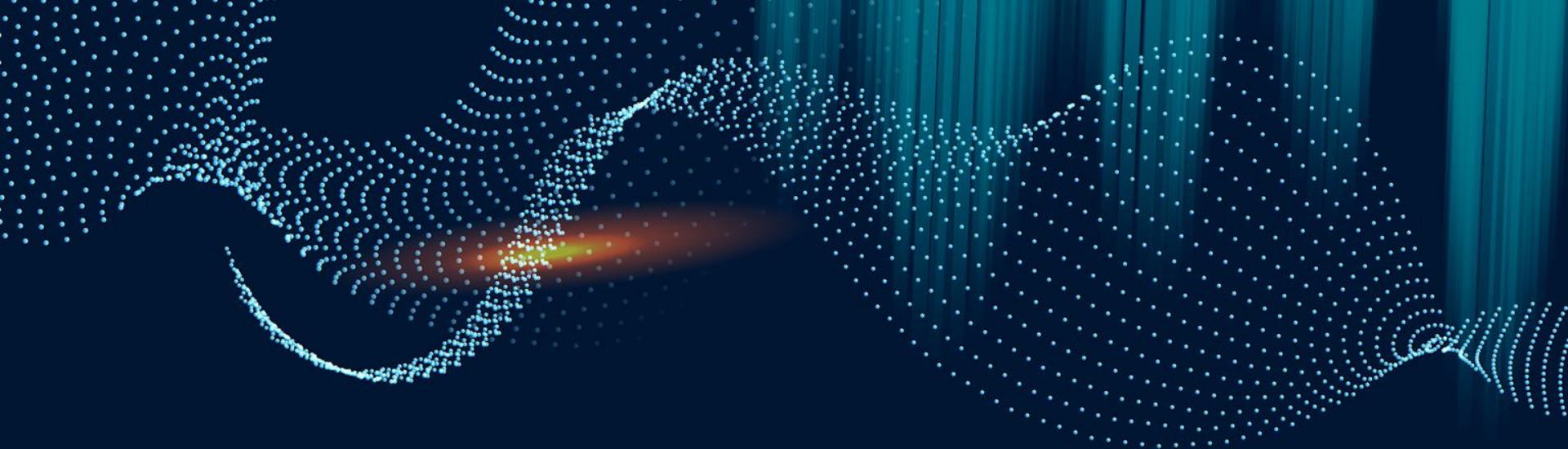
Terakhir (yeay), Saran sederhana tapi fatal

1. Buat timeline

- Tetapkan Langkah-Langkah besar
- Tentukan deadline untuk setiap Langkah
- Konsisten!

2. Buat log sheet

- Catat setiap hal yang dilakukan
- Jadikan deadline dari timeline sebagai acuan



1 Okt 2021

Data awal
sudah siap

21 Okt 2021

Atasi variansi
dulu



Harus sudah
tuntas pahami
paper

15 Okt 2021

Model
sederhana
sudah
dibentuk

31 Okt 2021

Contoh log sheet (versi 1)

Waktu	Yang dilakukan	Hasil	File luaran	Komentar
12 Des 2020	First run	0.65 train acc 0.6 val acc	start.csv	Underfit sepertinya
13 Des 2020	Perpanjang iterasi	0.75 train acc 0.72 val acc	Longer_epoch. csv	Well, better, tpi mungkin masih kurang
15 Des 2020	Perpanjang lagi iterasi	0.88 train acc 0.7 val acc	Longer_epoch_ 2.csv	Ups, malah overfit
16 Des 2020	Coba regularisasi	0.89 train acc 0.88 val acc	Regularization. csv	Nice. Tapi malah jadi agak berat
1 Jan 2021	Deadline 2: Bias dan variansi harus sudah kecil			

Contoh log sheet (versi 2)

Waktu	Epch	Batch	Reg	GPU	Arsi	file	Run time	Metrik	Komentar
12 Des 2020	50	8	No	No	CNN	1.Csv	30s	0.6, 0.65	Coba dulu arsi lain
13 Des 2020	50	8	No	No	LSTM	2.Csv	45s	0.77, 0.7	Sepertinya LSTM masih bisa di-tuning, tapi mau coba arsi lain dlu
15 Des 2020	100	8	No	No	BERT	3.Csv	120s	0.89, 0.8	Nah ini. Tapi terlalu kompleks
16 Des 2020	75	32	Yes	No	BERT	4.csv	100s	0.9, 0.9	Sekarang tuning-tuning biar perfect
1 Jan 2021	Deadline 2: Bias dan Variansi harus sudah kecil								

Terima Kasih

CREDITS: This presentation template was created by Slidesgo, including icons by Flaticon, and infographics & images by Freepik.

