

Kiat Sukses Proyek Data Analitik

Ministry of Finance - Data Analytics Community



A woman in a white hoodie is standing and pointing at a red flipchart in a meeting room. Several people are seated around a table, looking at the presentation. In the background, other people are standing near a large blue display. The room is dimly lit with warm ambient lighting.

Kenapa saya harus membuat proyek data analitik?

6 Langkah Memulai Proyek Data Analitik



1. Harus punya tujuan
2. Tentukan ekspektasi
3. Pahami data
4. Terjemahkan masalah
5. Rencanakan development & deployment
6. Tentukan *success metrics*

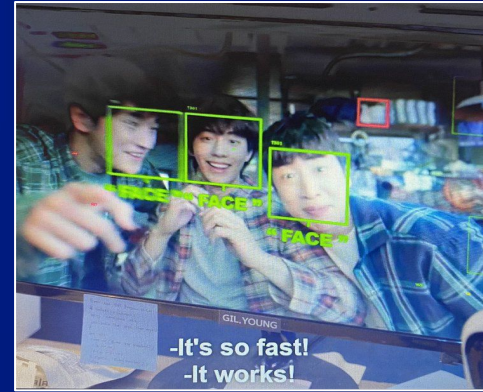
1 | Harus Punya Tujuan



Menyelesaikan masalah?



Melihat ada peluang?



Atau cuma ngikutin tren biar keren?

 You Retweeted



Ariya Hidayat  @AriyaHidayat · Jan 3

I wish more companies will realize that slapping any labels, buzzwords, or latest technology trends on their product won't magically make it amazing.



1



3



5



[Show this thread](#)

1 | Harus Punya Tujuan

- Masalah vs peluang
- Meningkatkan pendapatan vs mengurangi biaya
- Improvisasi vs inovasi
- Cost and time vs differentiation

Kita mungkin punya banyak tujuan... tapi



2 | Tentukan Ekspektasi

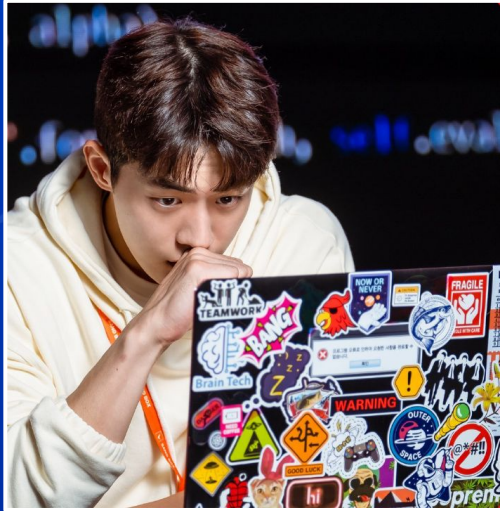


Apa anda sudah tahu *use case* yang ada di industri anda?



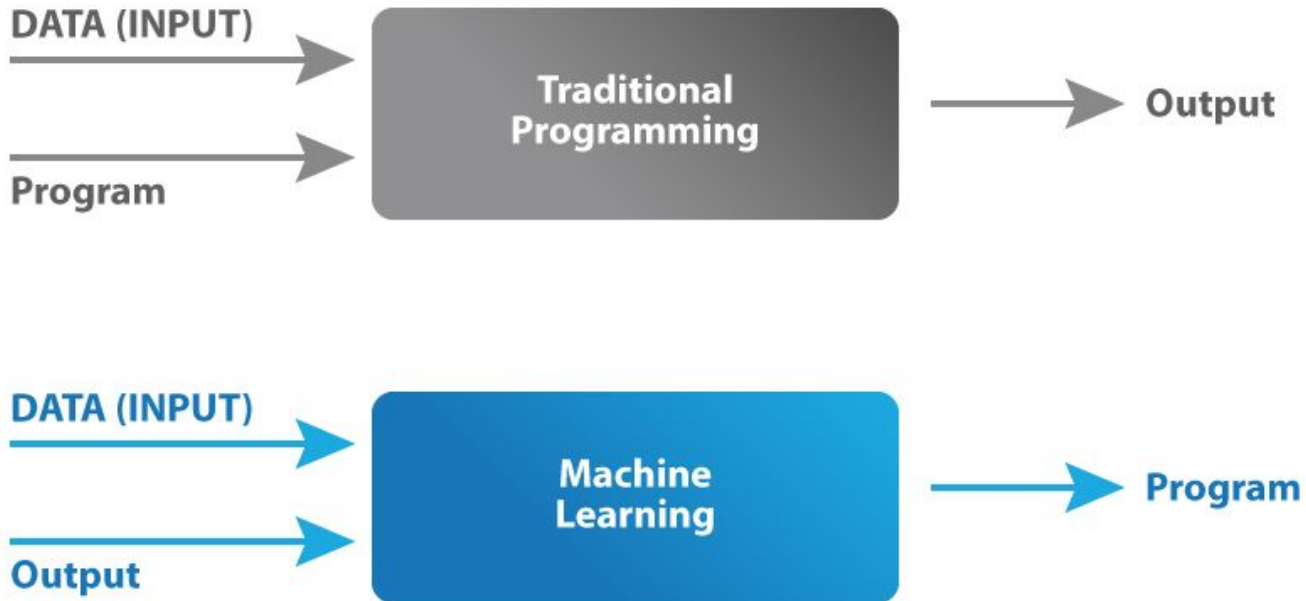
Bagaimana orang lain sebelumnya menyelesaikan masalah serupa?

2 | Tentukan Ekspektasi

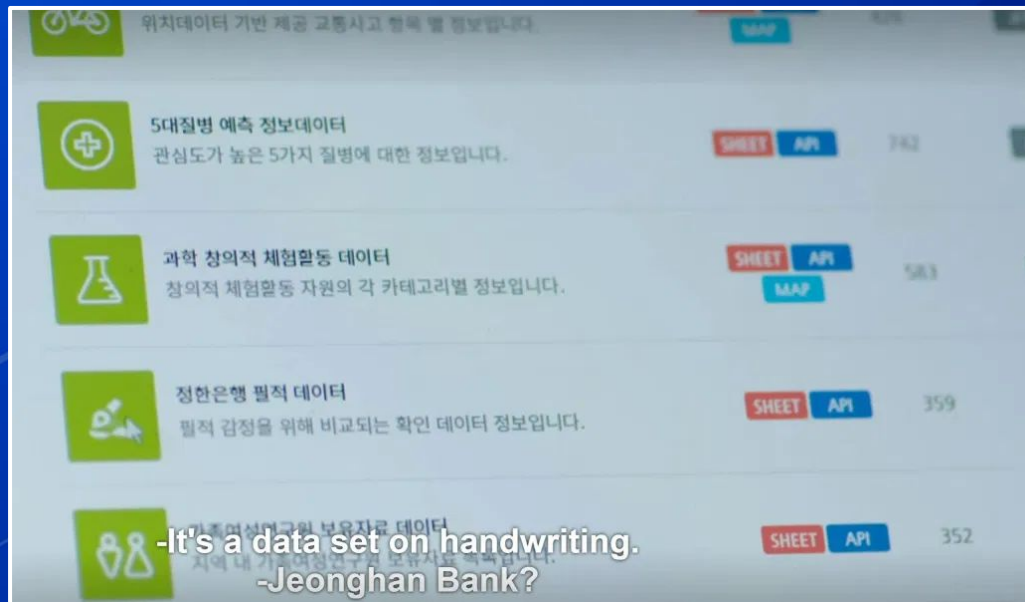


**Apa perlu data analitik?
Mungkin pendekatan
tradisional cukup.**

2 | Tentukan Ekspektasi

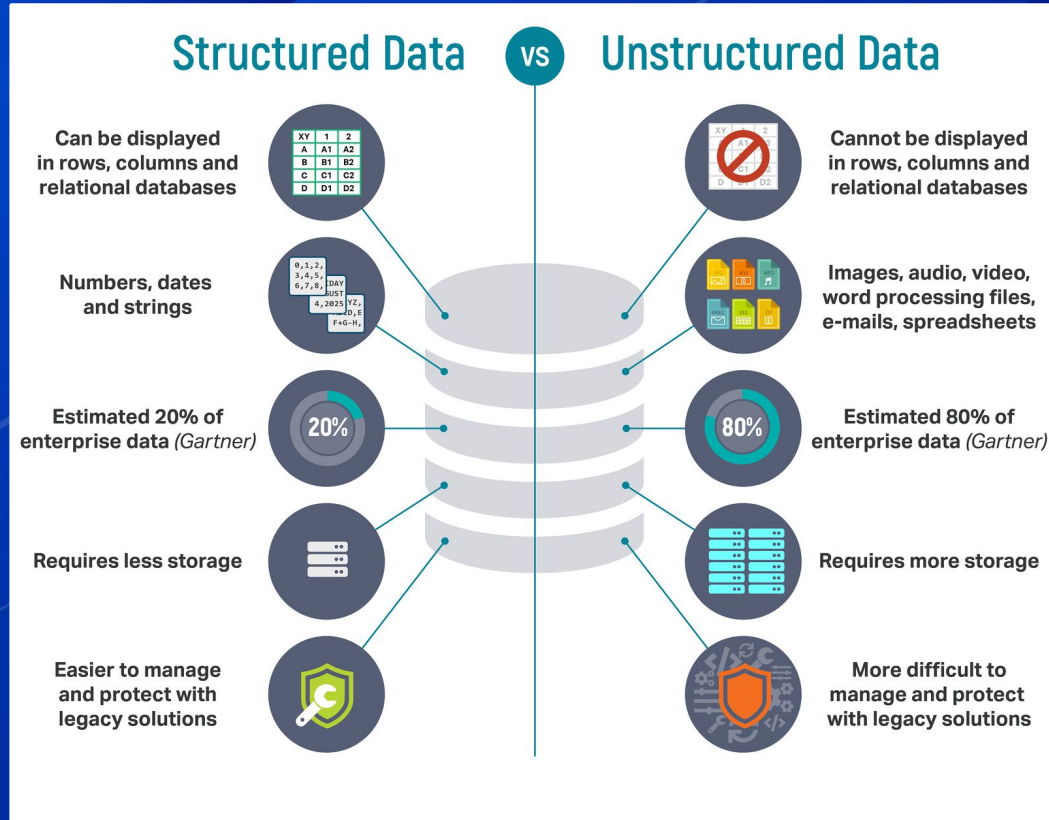


3 | Pahami Data



Datanya ada?

3 | Pahami Data



Pahami
jenis data

3 | Pahami Data

Jenis data akan mempengaruhi langkah pembuatan model pada proyek data analitik yang akan dilakukan



Jenis Data	Algoritma	Library
Structured	Logistic Regression, Decision Tree Family, GLM	Pandas, Apache Spark, Scikit-learn, XGBoost
Image	CNN (Inception, ResNet, ...)	Tensorflow, Keras, OpenCV
Text	RNN, Transformer (BERT, GPT, ...)	Tensorflow, Keras, Spacy, Gensim

3 | Pahami Data



Berapa
jumlahnya?
Apakah cukup?

3 | Pahami Data



Bagaimana
data disimpan?

3 | Pahami Data



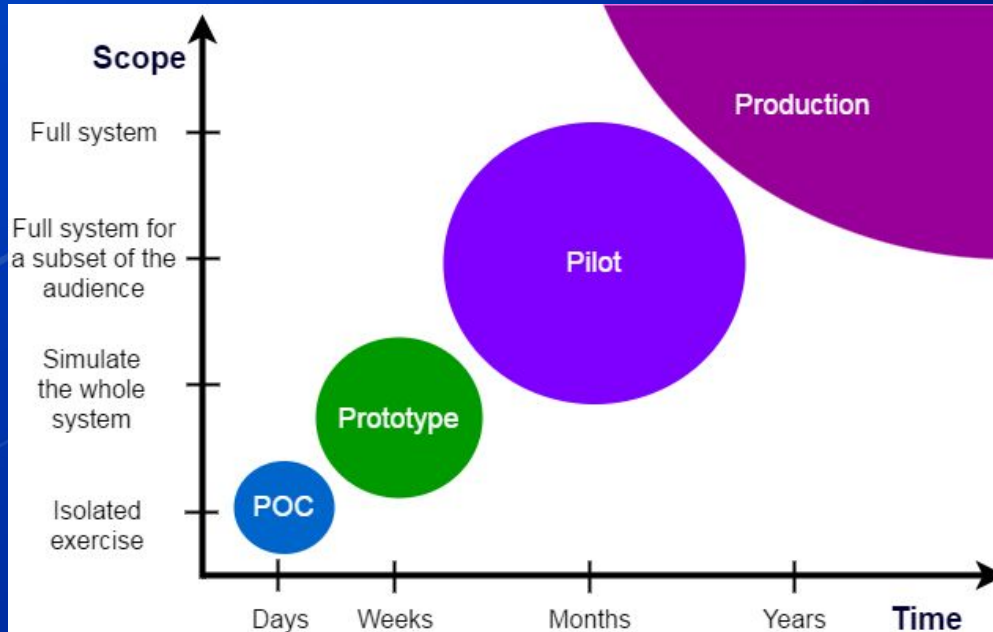
Bagaimana
kualitas
datanya?

3 | Terjemahkan Masalah Bisnis

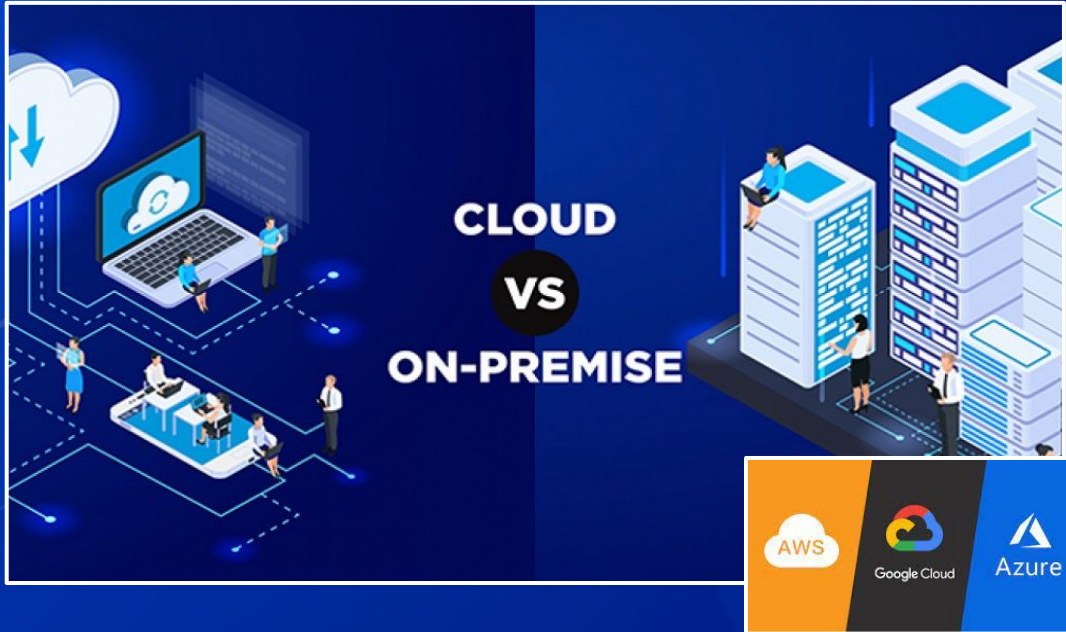
Masalah Bisnis	AI Problem
Skor Credit	Binary Classification
Klasifikasi Dokumen	Multiclass Classification
Deteksi Fraud	Binary Classification
Prediksi Pendapatan	Regression
Analisis Sentimen Sosial Media	NLP: Text Classification
Absensi Face Recognition	CompVis: Image/Face Classification

4 | Rencanakan Deployment dan Development

Proof of Concept (PoC) vs production



5 | Rencanakan Deployment dan Development



Pemilihan ini menentukan jenis library yang akan digunakan

6 | Tentukan *Success Metric*



Kita perlu mendefinisikan target sukses dari proyek data analitik yang akan kita kerjakan

Set Pola Pikir





Terus Belajar. Pantang Menyerah.

**Kelihatannya Data Analitik itu Susah.
Terlalu banyak istilah teknis yang saya gak tahu**



**Terakhir.
Kolaborasi itu
penting!**

Silakan Bookmark!

1. <https://github.com/datascienceid/datasciencecurriculum>
2. <https://github.com/khuyentran1401/awesome-Python-data-science-books>
3. <https://data-flair.training/blogs/data-science-tutorials-home/>
4. <https://github.com/eugeneyan/applied-ml>
5. <https://github.com/sastrawi/nlp-bahasa-indonesia>