



# FINAL ROUND



DATA SCIENCE BOOTCAMP BATCH

XVII



# BL OOM'S TAXONOMY



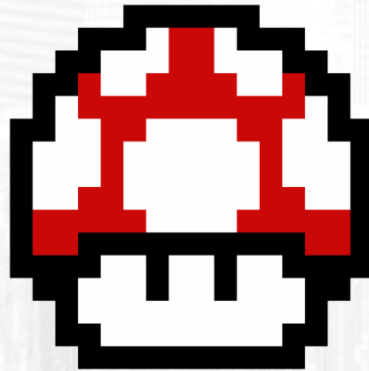
When someone learning by creating, they will earn much more experience than learning just to remembering

EXP +++

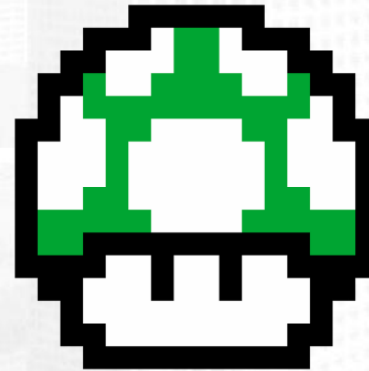


# DO NOT FORGET THE DATA SCIENCE SKILL-SET

*Practice, practice,  
and practice*



**PROGRAMMING AND  
COMPUTER SCIENCE**



**SUBJECT MATTER  
KNOWLEDGE**



**MATH AND  
STATISTICS**

*Learn, test, repeat!*

*Analyze many used case,  
read lot of book and article*

# FINAL PROJECT DESCRIPTION

YOU'LL TEAM UP WITH YOUR ALLIES, AND SOLVING DATA  
PROBLEMS BY IMPLEMENTING ALL MODULES YOU'VE  
LEARNED TO PRODUCE DATA-DRIVEN INSIGHTS



## STAGE 0 ---

Preparation

24<sup>th</sup> Jan - 30<sup>th</sup> Jan

## STAGE 1 ---

EDA, Insights &  
Visualization

31<sup>st</sup> Jan - 6<sup>th</sup> Feb

## STAGE 2 ---

Data Pre-  
Processing

7<sup>th</sup> Feb - 13<sup>rd</sup> Feb

## --- STAGE 3 ---

Machine Learning  
Modelling &  
Evaluation

14<sup>th</sup> Feb - 20<sup>th</sup> Feb

## --- STAGE 4 ---

Final  
Preparation

21<sup>st</sup> Feb - 27<sup>th</sup> Feb

## FINAL STAGE

25<sup>th</sup> Mar



SEE BOSS OPTIONS



CHECK HINTS!



# THE JUDGES!



RAKAMIN  
TEAM



TUTOR  
LEAD



INDUSTRY  
EXPERT



BUSINESS  
EXPERT





# IN ORDER TO PASS THE PROJECT YOU NEED TO DELIVER 4 DOCUMENT



## LAPORAN PROJECT

Berisi seluruh dokumentasi pengerjaan project dari tiap stage



## SOURCE CODE

Jupyter notebook yang berisi source code python, mulai dari eksplorasi data, pre-processing, EDA & penggalian insight, modeling, hingga evaluasi



## NOTULEN MENTORING

File PDF berisi rangkuman dari tiap sesi mentoring yang berlangsung



## MATERI PRESENTASI

File dalam format ppt untuk presentasi akhir

# LISTS OF EVALUATION

## KONTRIBUSI PEKERJAAN INDIVIDU – 40 %

Dinilai oleh mentor, yang dievaluasi berdasarkan:

(i). Keaktifan saat mentoring, (ii). Kontribusi yang diberikan dalam pengerjaan project

## LAPORAN AKHIR GRUP – 20 %

Dinilai oleh tutor lead (buat se detail mungkin), yang dievaluasi berdasarkan:

(i). Kejelasan isi (mudah dibaca dan dipahami) - 20% (ii). Kreativitas dalam menggali insight - 30%

(iii). Ketepatan teknik dalam eksperimen modeling (termasuk pre-processing) - 50%

## PRESENTASI AKHIR – 40 %

Dinilai oleh para juri, yang dievaluasi berdasarkan:

(i). Keselarasan antara problem yang disampaikan, dan bagaimana insight & model yang diberikan dapat membantu - 60%

(ii). Kejelasan saat penyampaian informasi - 15%

(iii). Kemampuan dalam menjawab pertanyaan - 20%

(iv). Ketepatan waktu presentasi (jika 10-15 menit, dapat nilai full, jika tidak, dapat 0) - 5%

# KNOWING YOUR PARTY AND CHOOSING THE BOSS!



## GRUP SUDAH DITENTUKAN SECARA ACAK

Kami sudah membuat grup yang anggotanya dipilih secara acak (dalam pekerjaan kamu tidak bisa memilih dengan siapa kamu bekerja), detail dapat dicek di lampiran



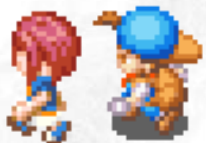
## TENTUKAN DATA-SET YANG AKAN DIOLAH

Silakan berdiskusi dan tentukan data yang akan diolah, data-set sudah tercantum di file ini dan akan diupload di LMS, silakan komentar detail kelompok di grup dan sebutkan data-set yang dipilih



## BAGI TUGAS DAN LANGSUNG HAJAR!

Jangan tunda untuk berdiskusi dan membagi tugas, karena penundaan beberapa hari akan sangat berpengaruh pada performa tim-mu





# **DO NOT GET LOST!**

## **PRESS THE HELP BUTTON AND SPEAK TO THE MENTOR**



Setiap grup wajib melakukan mentoring untuk lanjut ke stage berikutnya



Setiap anggota grup wajib hadir dalam sesi mentoring



Dalam sesi mentoring akan dilakukan penilaian terkait keaktifan individu, dan kualitas output pada stage tersebut



Sesi mentoring maksimal berlangsung selama 60 menit, tiap grup melaporkan progress sesuai format terlampir



Rangkuman sesi mentoring wajib diupload ke LMS sesuai nama grup yang tertera



# PERFECT MENTORING SCORE IS ACHIEVEABLE ! ! AS LONG AS



Aktif **bertanya** disaat mentoring. Pertanyaan yang dianggap sebagai poin adalah :

- a. Pertanyaan yang berkualitas : insightful, bersifat validatif, menunjukkan bahwa penanya sudah pernah mempelajari hal yang ditanyakan. Contoh : bukannya kalau hyperparameter tuning itu kita bisa gunain RandomizedSearchCV ya bang biar dapetin parameter terbaik? biar gak ngabisin waktu gitu?
- b. Pertanyaan bercabang dan terstruktur : pertanyaan berkualitas yang setelah dijawab, muncul pertanyaan lain sebagai respon atas jawaban tersebut sehingga terbentuk semacam diskusi berkualitas (pertanyaan berkualitas yang berkelanjutan)



# PERFECT MENTORING SCORE IS ACHIEVEABLE ! ! AS LONG AS



Aktif dalam memberikan masukan pengerjaan tugas baik kepada sesama student maupun mentor, terutama pada saat diberi kesempatan berkomentar (akan disediakan waktu khusus). Masukkan yang dianggap sebagai poin adalah :

- a. Masukan tunggal : masukan hanya terfokus terhadap masalah yang sedang dihadapi, tidak jauh kedepan setelah masalah tersebut diselesaikan. Contoh : "Kita coba gunain hyperparameter tuning aja"
- b. Masukan masif : masukan bersifat kompleks dan sistemik (seperti metode pengerjaan dengan langkah langkahnya), menunjukkan bahwa siswa memahami langkah yang harus dilakukan jauh kedepannya. Contoh : "Lebih baik sebelum ngelakuin penskalaan, datasetnya dipecah jadi train test dulu aja. Biar konstanta yang dipake pas fit itu gak mencakup data yang akan dipake sebagai test data. Abis itu diskalainnya pake MinMax Scalar, tapi kalau mau make Standard Scaler juga gak papa, datanya gak skewed kok"





# PERFECT MENTORING SCORE IS ACHIEVEABLE ! ! AS LONG AS



Mengerjakan tugasnya secara maksimal, sesuai dengan **komitmen** yang diberikan setiap minggunya



**Aktif** dalam berdiskusi pengerjaan tugas, dan berinisiatif dalam mengerjakan tugas, menawarkan pengerjaan tugas, atau mengusulkan pembagian tugas



Terlibat aktif dalam **setiap bagian pekerjaan**, termasuk coding, pembuatan Laporan dan PPT. Esensi dari tugas akhir adalah melatih kemampuan dalam Machine Learning, menyusun Laporan dan PPT yang informatif, insightful, dan deliverable. Maka setiap student wajib terlibat dalam semua bagian agar Learning Experiencenya maksimal



**Memahami apa** yang dikerjakan, **mengapa** mengerjakan dengan cara tersebut, dan memahami **alternatif** lain pengerjaannya apabila ada





# PROJECT AWARDS!



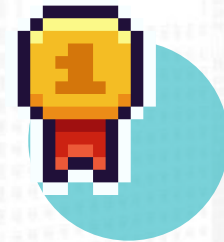
**MOST  
OUTSTANDING  
GROUP**

**+10 EXP**

Get highest score from  
mentoring session

**+20 EXP** ★

Get highest score  
from overall stages



**MOST  
OUTSTANDING  
INDIVIDUAL**

**+15 EXP** ★

Become a presenter in  
final presentation

**+2 EXP**

Posting stories or feed and  
tag @rakaminacademy

**TOTAL AWARDS WORTH RP5.MIO**

# *FINAL PROJECT* *GUIDANCE*

Check in LMS

# CHOOSE YOUR FINAL BOSS!

I mean your data set



## ECOMMERCE SHIPPING DATA:

Predict whether the Product Shipment Delivered on time or not

[See the Data](#)

# FINAL PROJECT PART IES

**1**

- Fakhri N
- Dyah Phitaloka
- Dwi M Nurafli
- Ryan Rizky F
- M Iqbal Tawakkal
- Awal kurnia achmad

**2**

- Satriahady A P
- M Annas
- Oktri Pani W
- Christopher S
- Defri Subastian
- Willi Bagus S

**3**

- TRI SETIAWAN
- Risa
- Ulva Dewiyanti
- Cristanto
- Steven Benny
- Tista Amelia

**4**

- Mira Ganis
- Keizia Duma
- Tri Pramudya
- Yohanes Bima
- Ikko Ananda L

**5**

- Mhd Fahmi Aziz
- Andre Yudha P
- Dwi Susanto
- Nur Ayu Asyifa
- Ryan Hardi

**6**

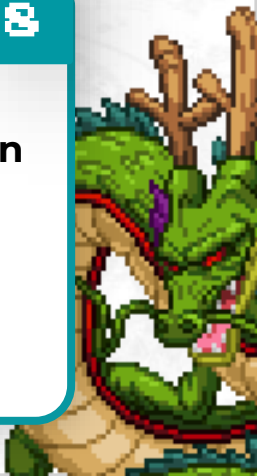
- Rima Nugrahiya
- Zaki kamal
- Rheza Fathur R
- Rachdi farisma
- Ali Rohmansyah

**7**

- Maysarah Ulfah
- Bryan Tamin
- Arif Rahman
- Vicken
- Wildan R R

**8**

- Ilham Ibnu A
- M Farhan A
- Fikri Diva S
- Dharma Setiawan
- Ahmad Ilham H





# Initial Execution



Perwakilan tiap kelompok mengirimkan ke Class Coordinator:

- Nama Kelompok
- Ketua Kelompok
- Pilihan dataset



Buat grup diskusi kelompok (ex : WA), undang Mentor ke grup diskusi, connect dengan Linked.in mentor , mulai mengerjakan Stage 0



# First Mentoring Execution

## First step



Tentukan komitmen pengerjaan untuk Stage 1



Satukan komitmen minggu ini, dan berikan ke Mentor, dan tentukan jadwal mentoring



## Gap time

Masing-masing anggota mengerjakan komitmennya



## Live

Mentoring pertama (komitmen dicek, dan buat komitmen untuk Stage selanjutnya)



# STAGE 0

## Preparation



### PROBLEM STATEMENT

Sudah menentukan problem apa yang akan diselesaikan dengan model dari dataset yang dipilih.



### ROLE

Sudah memahami peran yang akan dijalankan dalam mengolah dataset.



### GOAL

Sudah menentukan tujuan utama yang ingin dicapai.



### OBJECTIVE

Sudah menentukan sasaran akhir yang dicapai dari suatu tujuan yang telah ada.



### BUSINESS METRICS

Sudah mengembangkan *business metrics* untuk mengukur ketercapaian objective yang sudah ditentukan.

BACK TO HOME





# STAGE 1

## EDA, Insights & Visualization



### EXPLORATORY DATA ANALYSIS

Untuk setiap feature yang disiapkan sudah dicek distribusinya terhadap variabel target



### INSIGHTS AND VISUALIZATION

Sudah punya beberapa insight menarik yang diperoleh dari dataset, beserta visualisasi yang mendukung. Insight dikaitkan dengan masalah utama yang ingin diselesaikan

**GAME  
OVER!**

### COMMON PITFALLS

Insight yang disajikan tidak memiliki relevansi pada masalah utama, atau tidak mengaitkan temuan pada sumber lain di luar data



BACK TO HOME





# STAGE 2

## Data Pre-Processing



### DATA EXPLORATION

Dataset sudah di-unduh dan di-explore apa saja attributes-nya dan targetnya output olahan datanya



### DATA CLEANSING

Sudah melakukan pengecekan data bermasalah (missing values, invalid values, data duplicates), dan sudah membersihkannya



### FEATURE ENGINEERING

Sudah menentukan feature apa saja yang akan digunakan, feature apa yang perlu ditambahkan, dan reformatting feature sesuai kebutuhan



BACK TO HOME



# STAGE 3

## Machine Learning Modelling & Evaluation



### MODELLING

Membuat beberapa modelling Machine Learning sesuai dengan objektif yang ingin dicapai



### EVALUATION

Melakukan evaluasi dari beberapa model yang dibuat, dan memilih evaluasi dengan hasil terbaik

**GAME  
OVER!**

### COMMON PITFALLS

Tidak banyak eksperimen dalam feature engineering, terlalu fokus dalam mencoba-coba algoritma model & parameter, kurang mendalam dalam evaluasi model



BACK TO HOME



# STAGE 4

## Final Preparation



### PRESENTATION MATERIALS

Memosisikan diri untuk mempresentasikan hasil pengolahan kepada user dari divisi bisnis. Pastikan struktur materi yang disampaikan mudah dinikmati, runut, dan to-the-point



### SPEECH PRACTICE

Mempersiapkan diri untuk mempresentasikan hasil akhir kepada dewan juri, pada tahap ini grup berlatih di depan mentor dan tim Rakamin

**GAME  
OVER!**

### COMMON PITFALLS

Terlalu teknis dalam mempresentasikan hasil pengolahan data, visualisasi kurang mudah dipahami, tidak dapat menjelaskan dampak bisnis dari keputusan yang direkomendasikan



BACK TO HOME



# FINAL STAGE!



## RESULT PRESENTATION

Setiap grup diberikan waktu 15 menit presentasi dan 15 menit tanya jawab, grup boleh memilih 1 presenter atau lebih

**hint:** Juri adalah orang yang sangat ahli di bidangnya masing-masing



## ANSWERING JUDGES

Juri akan memberikan pertanyaan bagi peserta, dan jawaban tidak boleh diberikan oleh presenter (kecuali jika semua anggota melakukan presentasi)

**GAME  
OVER!**

## COMMON PITFALLS

Terlalu teknis dalam mempresentasikan hasil pengolahan data, visualisasi kurang mudah dipahami, tidak dapat menjelaskan dampak bisnis dari keputusan yang direkomendasikan



BACK TO HOME





# PROJECT HINTS!

## ○ START WITH WHY

**Sebelum memulai memproses data, tentukan terlebih dahulu masalah yang ada di lapangan dengan mencari berita atau kasus yang relevan, dan tentukan objektif utama yang ingin dicapai**

Contoh: Sebuah koperasi simpan pinjam masih melakukan assessment kelayakan kredit secara manual, hal ini membutuhkan waktu yang lama dan memperbesar kemungkinan human-error. Sehingga dibutuhkan model untuk menilai kelayakan kredit secara otomatis

## ○ MANAGE YOUR TASK LISTS EFFECTIVELY

**Bagi tugas dengan baik, saling pantau progress, saling membantu, dan gunakan waktu sebaik mungkin, jangan biarkan tugas menumpuk di minggu-minggu akhir**

## ○ ASK THE MENTOR!

**Yup, setiap grup akan memiliki mentor, tanyakan hal-hal yang strategis, konsultasikan progresmu, pastikan bahwa kamu ada di track yang benar**



# LAMP IRAN

CEK DENGAN TELIT I SETIAP BAGIAN

# TEMPLATE LAPORAN

[Unduh Template](#)

# TEMPLATE NOT UL EN MENT ORING

[Unduh Template](#)





Nama Kelompok:  
Stage:  
Mentor:

**Pembagian tugas di stage ini:**

Nama:

Nama:

Nama:

**Poin pembahasan:**

1.

2.

3.

4.

Dst.

**Hasil Diskusi:**

**Tindak Lanjut:**

The background of the slide is a faded, grayscale image of a dense urban skyline with numerous skyscrapers and buildings.

# **TEMPLATE PRESENTASI**

**Tidak ada template khusus, justru di situ tantangannya. Pastikan bahwa slide yang kamu buat mencakup informasi berikut;**

- 1. Problem apa yang ingin diselesaikan?**
- 2. Dataset seperti apa yang dimiliki?**
- 3. Insight apa saja yang ditemukan dari data?**

Bisa dijelaskan top 2 insights yang paling keren, dan tidak hanya dijelaskan insightnya apa, tetapi juga **action** apa yang dapat dilakukan setelah mengetahui insight tersebut, karena tugas dari seorang data scientist tidak cukup hanya memberikan insight namun harapannya insight tersebut dapat secara konkrit diwujudkan dalam bentuk action yang jelas bagi perusahaan.

Contoh: PNS punya peluang lebih tinggi untuk diterima pengajuannya, yakni 75% pengajuan diterima. Tetapi secara keseluruhan, hanya sekitar 12% pengajuan yang dari PNS. Maka kita perlu membuat campaign agar lebih banyak PNS yang tertarik untuk mengajukan pinjaman.

- 4. Apa saja yang telah dilakukan dalam membuat model**

Tidak perlu dijelaskan eksperimen yang sudah dilakukan apa saja, lebih fokus ke model final yang diperoleh. Itu pakai features apa saja, pre-processingnya bagaimana, pakai algoritma apa, dsb. Lalu, performance nya bagaimana.

**Perbanyak referensi agar pemahamanmu lebih kaya, dan dapat memahami usermu dengan lebih baik**