

# Financial Data Challenge 2022



# Financial Data Challenge

Description

Evaluation

Data

Rules

## LATAR BELAKANG:

Challenge ini dibuat atas kerjasama PT. Sharing Vision Indonesia dengan BRI (PT Bank Rakyat Indonesia, Tbk) dengan tujuan:

- Membuka kesempatan untuk peserta dari seluruh Indonesia, dari Sabang sampai Merauke untuk merasakan Challenge terhadap real financial use case.
- Membuka peluang peserta untuk mendapatkan kesempatan bekerja di BRI/Sharing Vision.

## USECASE:

Goal: Membuat model klasifikasi yang dapat menentukan apakah seorang nasabah akan **churn atau tidak**.

Model harus dibuat sekurang-kurangnya meliputi :

1. *Exploratory data analysis (EDA)*
2. *Feature Engineering*
3. *Modeling*
4. *Model Evaluation*

# Financial Data Challenge

Description

Evaluation

Data

Rules

## EVALUASI DALAM KOMPETISI:

Evaluasi dilakukan dengan mengukur performa model yang dilakukan oleh peserta, berdasarkan performance AUC yang dihasilkan dari data testing.

Peserta diwajibkan mengirimkan file melalui link: [bit.ly/Result\\_FDC2022](https://bit.ly/Result_FDC2022) dengan items sebagai berikut:

1. Coding script dalam bentuk '**nama\_lengkap-submission.ipynb**'
2. Nama file output (hasil prediksi terhadap data back testing): '**nama\_lengkap-submission.csv**', yang terdiri dari 2 (dua) kolom dengan nama kolom: 'Id' dan 'Predicted'.

Mohon perhatikan bahwa nama file dan nama kolom bersifat case sensitive.

fin\_data\_challenge\_train.csv

fin\_data\_challenge\_test.csv

'nama\_lengkap-submission.ipynb'

'nama\_lengkap-submission.csv'

Coding script .ipynb sekurang-kurangnya mengandung data preparation, EDA, feature engineering, modeling, evaluation, output akhir hasil prediksi dari data back testing.

# Financial Data Challenge

Description

Evaluation

Data

Rules

Data terdiri atas dua dataset, yaitu:

- fin\_data\_challenge\_train.csv: 100.000 baris dan 126 kolom (125 kolom fitur dan 1 kolom target 'y')
- fin\_data\_challenge\_test.csv: 25.000 baris dan 125 kolom (tanpa kolom target 'y')

## Deskripsi Kolom

Data Nasabah:

x0-x124 (125 kolom) : Merupakan data nasabah yang telah dinormalisasi dan dirahasiakan nama kolomnya.

Variabel target:

y - Apakah nasabah akan churn? (1 :iya, 0 :tidak)

Note : terdapat beberapa *Missing Value* pada data.

# Financial Data Challenge

Description

Evaluation

Data

Rules

- Satu akun per peserta
- Tidak ada penggabungan peserta dalam kompetisi ini
- Batasan/larangan dalam kompetisi ini adalah sebagai berikut:
  - Peserta dilarang untuk melakukan code sharing
  - Peserta dilarang untuk meminta atau mendapatkan bantuan dari pihak lain
  - Peserta dilarang melakukan pelabelan manual

## Jadwal:

10 Feb 2022	: Webinar Roadmap to Smart Banking Mindset Through AI-ML
14 Feb – 11 Maret 2022	: Pelaksanaan Lomba
14 Mar – 18 Mar 2022	: Penilaian
25 Mar 2022	: Pengumuman Pemenang (KEPUTUSAN JURI TIDAK DAPAT DIGANGGU-GUGAT)

Note : Link dataset yang digunakan dalam lomba ini akan dikirimkan pada tanggal 14 Februari 2022

Terimakasih  
Danke Thank you  
Gracias Grazie Merci  
Arigatou Obrigado  
Syukron Gamsa Hamnida  
Xie-Xie Kheilil Mamnun

