

## Goals:

o Memprediksi gaji seseorang berada di atas 5 Juta atau dibawah 5 Juta

## Dataset

- Dataset diambil dari <a href="https://www.kaggle.com/c/sanbercode-data-science/data">https://www.kaggle.com/c/sanbercode-data-science/data</a>
- Terdiri dari 13 kolom dan 38393 baris
- 7 kolom digunakan untuk memprediksi fitur target
- Kolom yang digunakan ialah:
  - Umur
  - Pendidikan
  - Status Nikah
  - Jenis Kelamin
  - Jam Kerja Mingguan
  - Kelas Pekerja
  - Pekerjaan

## Modeling

- Menggunakan 4 algoritma machine learning
  - Linear Regression
  - KNN
  - Random Forest
  - XGBOOST

## Evaluasi model

• Model yang menunjukan performa paling tinggi ialah XGBOOST dari segi akurasi hingga F1-Score