

# Praktikum Machine Learning

## Bagian 1. Preprocessing dan Reduksi Dimensi

Maret 2022

Pada praktikum ini, pengerjaan dapat dilakukan hingga pukul 23.59 sehari sebelum praktikum selanjutnya. Dapat menggunakan bahasa pemrograman Python atau R. Jika ada pertanyaan dapat menghubungi asisten di MS Teams.

### 1 Import Data dan Preprocessing

Dengan menggunakan pandas, muatlah data dari file **melb\_data.csv** dan lakukan proses preprocessing seperti pada contoh latihan sebelumnya. Yang dapat dilakukan adalah:

1. Imputation
2. One hot encoding tidak harus, namun boleh kalau diperlukan
3. Thresholding dan binarizing tidak harus, namun boleh kalau diperlukan
4. Scaling dan normalizing
5. Bonus - Oversampling atau undersampling

### 2 Reduksi Dimensi—PCA

*Principal Component Analysis* (PCA) adalah metode reduksi dimensi dengan mengubah observasi yang dependen menjadi komponen yang independen. Setelah anda melakukan preprocessing untuk data **melb\_data.csv**, ekstraklah data menjadi 2 *principal component*. Lakukan machine learning untuk regresi dengan target variabel —price— dan amati perbandingan loss value sebelum dan setelah diekstrak menjadi komponen utama.