

Feature Engineering & Machine Learning untuk STLF

Teknik Sliding Window, Fitur Lag/Kalender, XGBoost, dan Validasi Silang Temporal

Ir. Novalio Daratha, S.T., M.Sc., Ph.D.

Program Studi Teknik Elektro – Fakultas Teknik
Universitas Bengkulu

Minggu 10 / Semester Ganjil 2026

Agenda Perkuliahan Minggu ke-10

Capaian Pembelajaran (Sub-CPMK 9)

Mahasiswa mampu merumuskan problem peramalan deret waktu ke dalam kerangka *Supervised Learning*, merancang rekayasa fitur temporal, serta melatih dan memvalidasi model *Gradient Boosting* (XGBoost/LightGBM).

- 1 Formulasi Supervised Learning
- 2 Teknik Rekayasa Fitur Temporal
- 3 Encoding Kalender Siklis (Sin/Cos)
- 4 Algoritma Tree-Based: XGBoost & LightGBM
- 5 Time Series Cross-Validation (Mencegah Data Leakage)
- 6 Feature Importance & Analisis Hasil

Mengubah Deret Waktu Menjadi Supervised Learning

Menggunakan teknik **Sliding Window / Lagged Matrix**:

Representasi Matriks Fitur (X) dan Target (y)

$$\begin{bmatrix} y_t \\ y_{t+1} \\ \vdots \\ y_N \end{bmatrix} = f \left(\begin{bmatrix} y_{t-1} & y_{t-2} & \dots & y_{t-p} & X_{\text{kalender},t} & X_{\text{cuaca},t} \\ y_t & y_{t-1} & \dots & y_{t-p+1} & X_{\text{kalender},t+1} & X_{\text{cuaca},t+1} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ y_{N-1} & y_{N-2} & \dots & y_{N-p} & X_{\text{kalender},N} & X_{\text{cuaca},N} \end{bmatrix} \right) \quad (1)$$

Keunggulan: Dapat menyerap puluhan fitur non-linier secara simultan tanpa asumsi distribusi kaku.

Taksonomi Rekayasa Fitur (*Feature Engineering*)

1. Fitur Lag Historis

- Lag dekat: $t - 1, t - 2, t - 3$
- Lag harian: $t - 24, t - 48$
- Lag mingguan: $t - 168$

2. Statistik Bergerak

- Rolling Mean (24 jam)
- Rolling Std Dev
- Rolling Max / Min
- Exponential Weighted MA

3. Fitur Kalender

- Jam dalam sehari (0–23)
- Hari dalam seminggu (0–6)
- Hari Libur Nasional (Flag biner 0/1)

Encoding Siklik (Jam & Bulan)

Nilai jam 23 dan jam 00 berdekatan secara fisik. Gunakan transformasi trigonometri:

$$x_{\sin} = \sin(2\pi \cdot \text{jam}/24) \text{ dan } x_{\cos} = \cos(2\pi \cdot \text{jam}/24).$$

Mencegah Data Leakage: Time Series Cross-Validation

Bahaya K -Fold CV Acak Standar

Pengacakan baris (*shuffling*) menyebabkan data masa depan bocor ke masa lalu (*Lookahead Bias / Leakage*). Model tampak memiliki akurasi 99% di lab namun gagal total di lapangan.

Metode Validasi yang Benar

Expanding Window (Time Series Split):

- Fold 1: Train [Bulan 1–3] → Test [Bulan 4]
- Fold 2: Train [Bulan 1–4] → Test [Bulan 5]
- Fold 3: Train [Bulan 1–5] → Test [Bulan 6]

Implementasi Python: XGBoost untuk STLF

```
import xgboost as xgb
from sklearn.metrics import mean_squared_error

# 1. Pipeline Feature Engineering
df['hour_sin'] = np.sin(2 * np.pi * df.index.hour / 24.0)
df['hour_cos'] = np.cos(2 * np.pi * df.index.hour / 24.0)
df['lag_24'] = df['load'].shift(24)
df['roll_mean_24'] = df['load'].shift(1).rolling(24).mean()
data = df.dropna()

# 2. Split Berurutan Temporal
X = data[['hour_sin', 'hour_cos', 'lag_24', 'roll_mean_24', 'temp']]
y = data['load']
X_train, X_test = X[:-168], X[-168:]
y_train, y_test = y[:-168], y[-168:]

# 3. Training Gradient Boosting
model = xgb.XGBRegressor(n_estimators=1000, learning_rate=0.02, max_depth=5, subsample=0.8)
model.fit(X_train, y_train, eval_set=[(X_test, y_test)], early_stopping_rounds=50, verbose=False)

# 4. Feature Importance
xgb.plot_importance(model, max_num_features=5)
```

Rangkuman & Persiapan Minggu ke-11

Poin Kunci Minggu 10

- Rekayasa fitur (lag, rolling, calendar encoding) adalah kunci akurasi algoritma Machine Learning.
- XGBoost dan LightGBM sangat tangguh menangani relasi non-linier kompleks beban sistem tenaga.
- Validasi temporal yang ketat tanpa shuffling menjamin keandalan model saat diterapkan pada sistem online.

Materi Minggu Depan (Minggu ke-11)

Deep Learning Dasar untuk Deret Waktu: Arsitektur Recurrent Neural Network (RNN), Long Short-Term Memory (LSTM), dan Gated Recurrent Unit (GRU) untuk data pembangkitan EBT.

Terima Kasih

Pertanyaan & Diskusi Machine Learning untuk Deret Waktu

Laboratorium Sistem Tenaga & Komputasi Cerdas
Jurusan Teknik Elektro – Universitas Bengkulu