

DAILY TEMPERATURE FORECASTING ANALYSIS FOR WEDDING PLANNERS TO OPTIMIZE EVENT PLANNING

Fika Nabila
JCDS 2804-019

CONTEXT

Perubahan iklim menyebabkan variabilitas suhu yang memengaruhi perencanaan acara pernikahan, khususnya di lokasi outdoor.

Anita Weddings, sebuah Wedding Organizer skala menengah di Jakarta, menghadapi tantangan dalam merencanakan acara pernikahan indoor dan outdoor akibat variabilitas suhu.

Dengan data iklim dari API, Anita Weddings ingin membangun model prediksi suhu harian untuk mendukung pengambilan keputusan lokasi dan kebutuhan perlengkapan acara.

PROBLEM STATEMENT

Bagaimana memprediksi suhu harian agar Anita Weddings dapat menentukan lokasi acara (indoor/outdoor) dan perlengkapan yang sesuai?

GOALS

- Memprediksi suhu harian untuk mendukung keputusan lokasi.
- Mengoptimalkan perlengkapan seperti tenda, kipas, atau pemanas.
- Meminimalkan risiko akibat perubahan suhu mendadak.

ANALYTIC APPROACH

Menggunakan data iklim (kelembaban, angin, tekanan udara, dll.) untuk membangun model regresi prediksi suhu.

METRIC EVALUATION

- RMSE & MAE : Mengukur rata-rata kesalahan prediksi.
- R^2 : Menilai seberapa baik model menjelaskan variansi suhu.

Semakin kecil RMSE/MAE dan semakin tinggi R^2 , semakin baik performa model.

FITUR-FITUR DALAM DATA

Fitur	Tipe Data	Deskripsi
city	object	Nama kota tempat data cuaca direkam
datetime	object	Waktu dan tanggal pengambilan data
temp	float64	Suhu udara saat observasi, dalam derajat Celcius
humidity	int64	Persentase kelembaban udara (0-100%)
pressure	int64	Tekanan udara dalam satuan hPa (hectopascal)
wind_speed	float64	Kecepatan angin dalam meter per detik (m/s)
weather_desc	object	Deskripsi kondisi cuaca (misal: "clear sky", "rain", dll.)

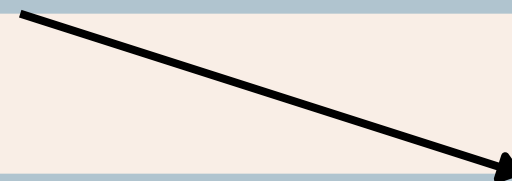
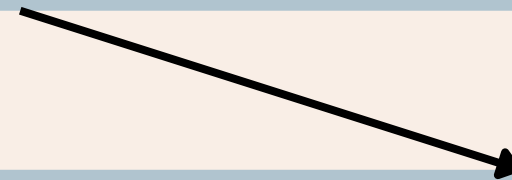
LOAD DATA

EKSTRAK DATA DARI OPENWEATHERAPI

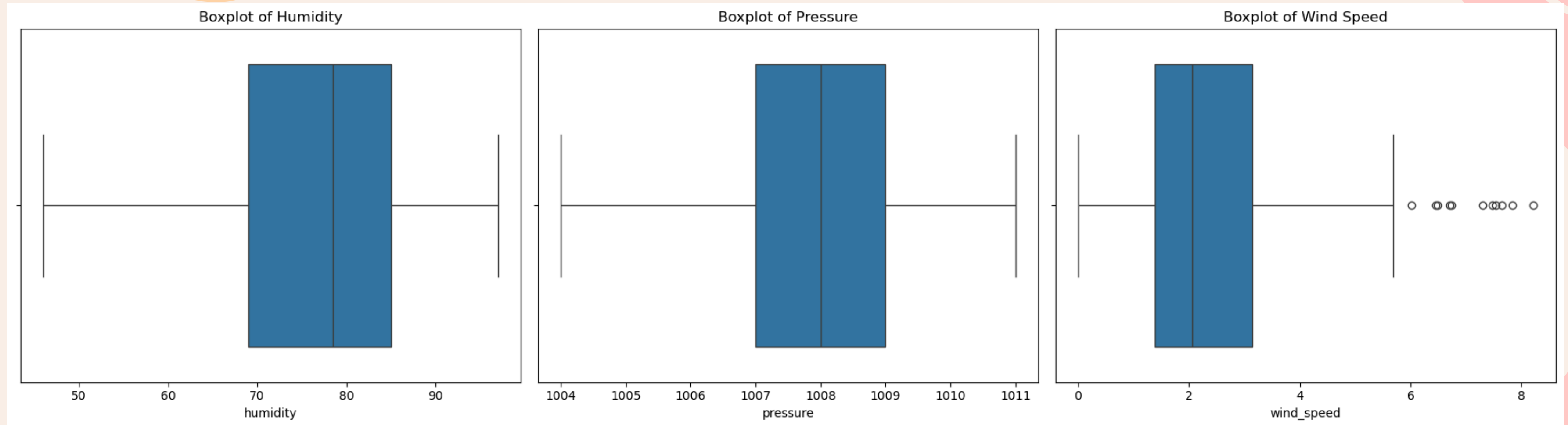
UBAH DATETIME KE TYPE DATA DATETIME

TAMBAH FITUR HOUR DAN WEEKDAY

SIMPAN KE CSV DAN LOAD ULANG



DATA OUTLIER



- Scaling dengan RobustScaler
- Outlier tidak dihapus karena masih dalam batas yang masuk akal, juga agar model tetap bisa memprediksi kasus ekstrem
- Transformasi fitur waktu menjadi fitur siklikal dengan sin dan cos (membantu model memahami pola waktu yang periodik – harian atau mingguan)

DATA PREPARATION

FITUR & TARGET

Fitur : ['city', 'humidity', 'pressure', 'wind_speed',
'weather_desc', 'hour', 'weekday']

Target : ['temp']

TRAIN TEST SPLIT

80%
Data Train

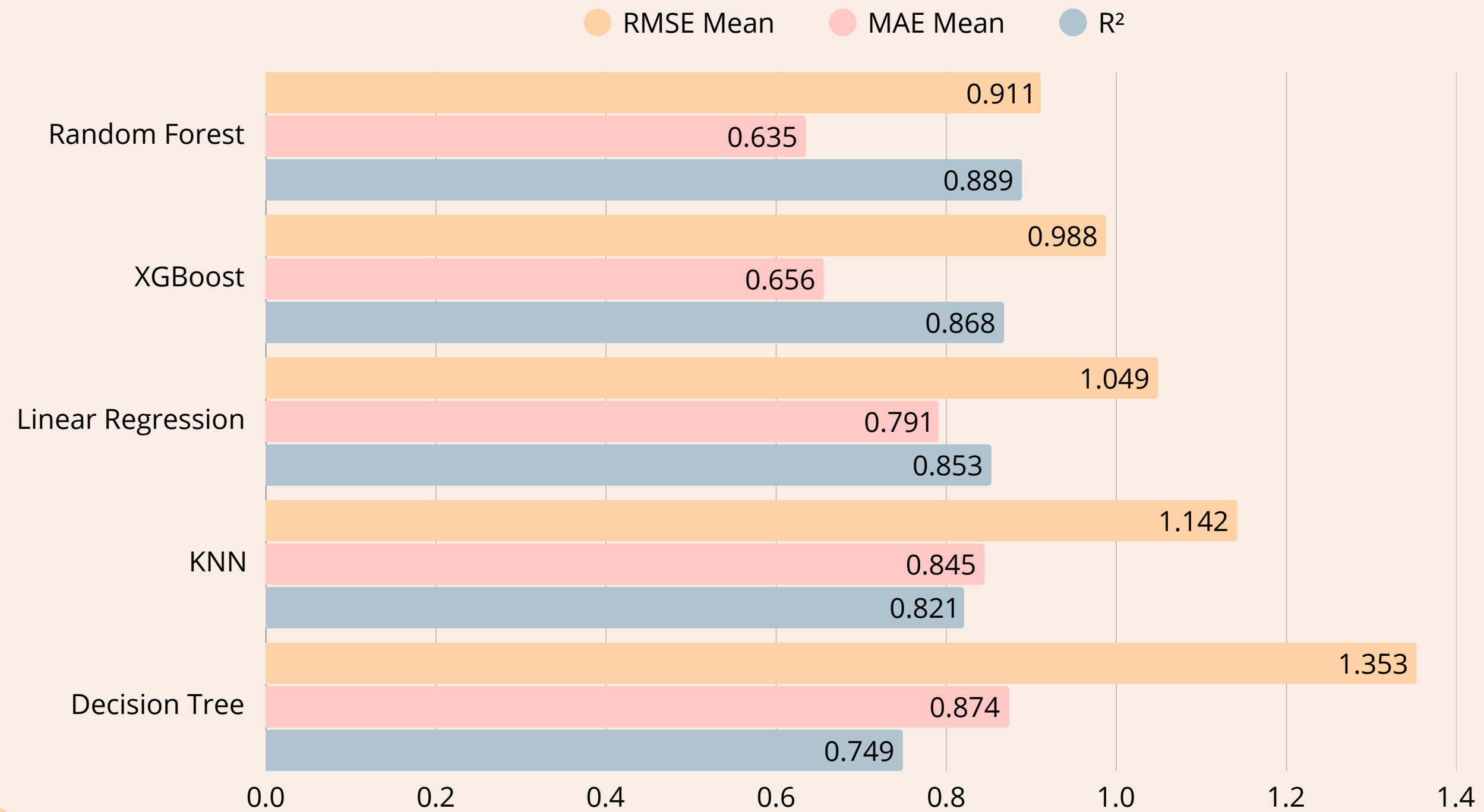
20%
Data Test

PREPROCESSING WITH PIPELINE

ColumnTransformer :

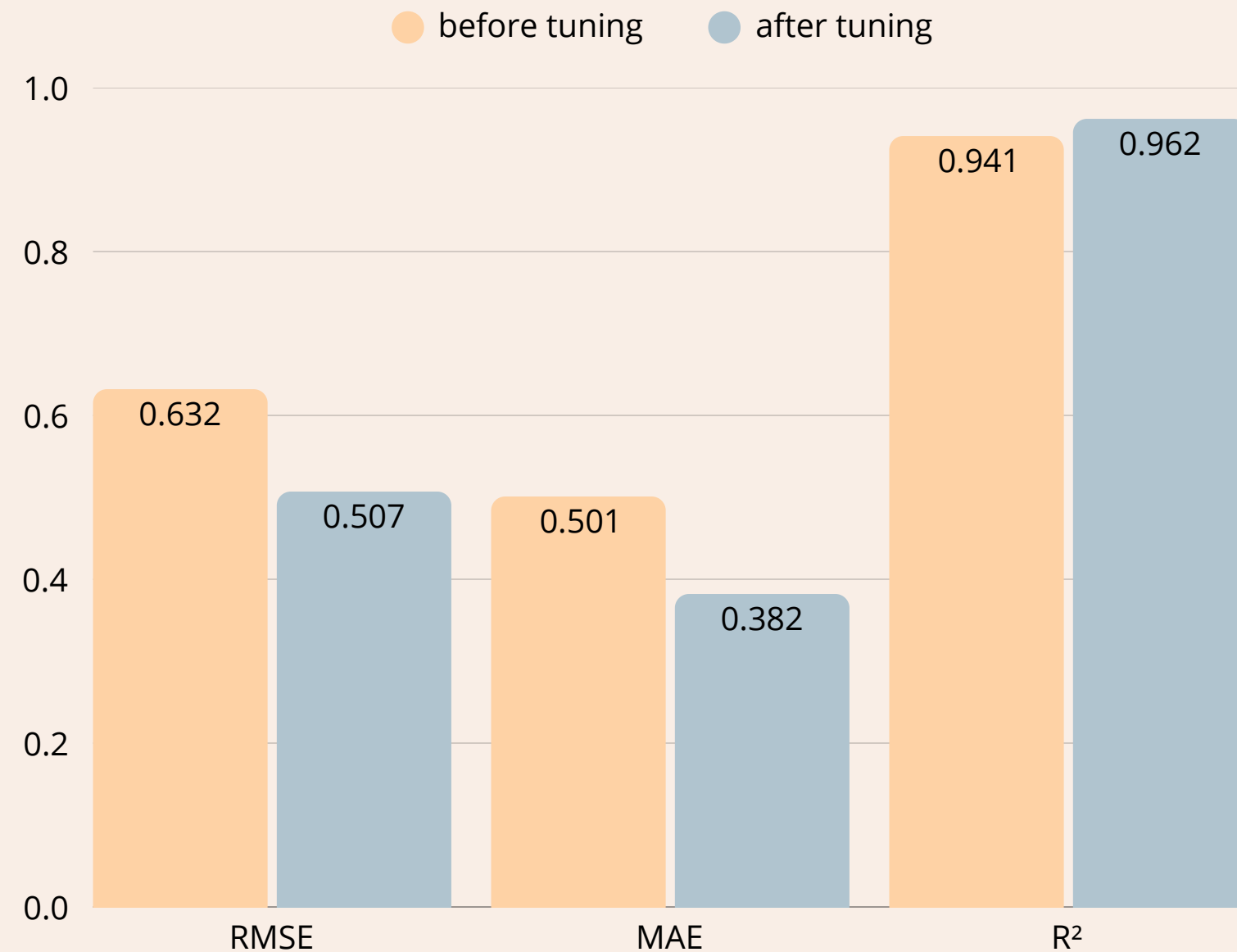
- time features → CyclicalFeatures()
- number features → RobustScaler()
- categorical features → OneHotEncoder()

DATA TRAIN : MODEL BENCHMARKING



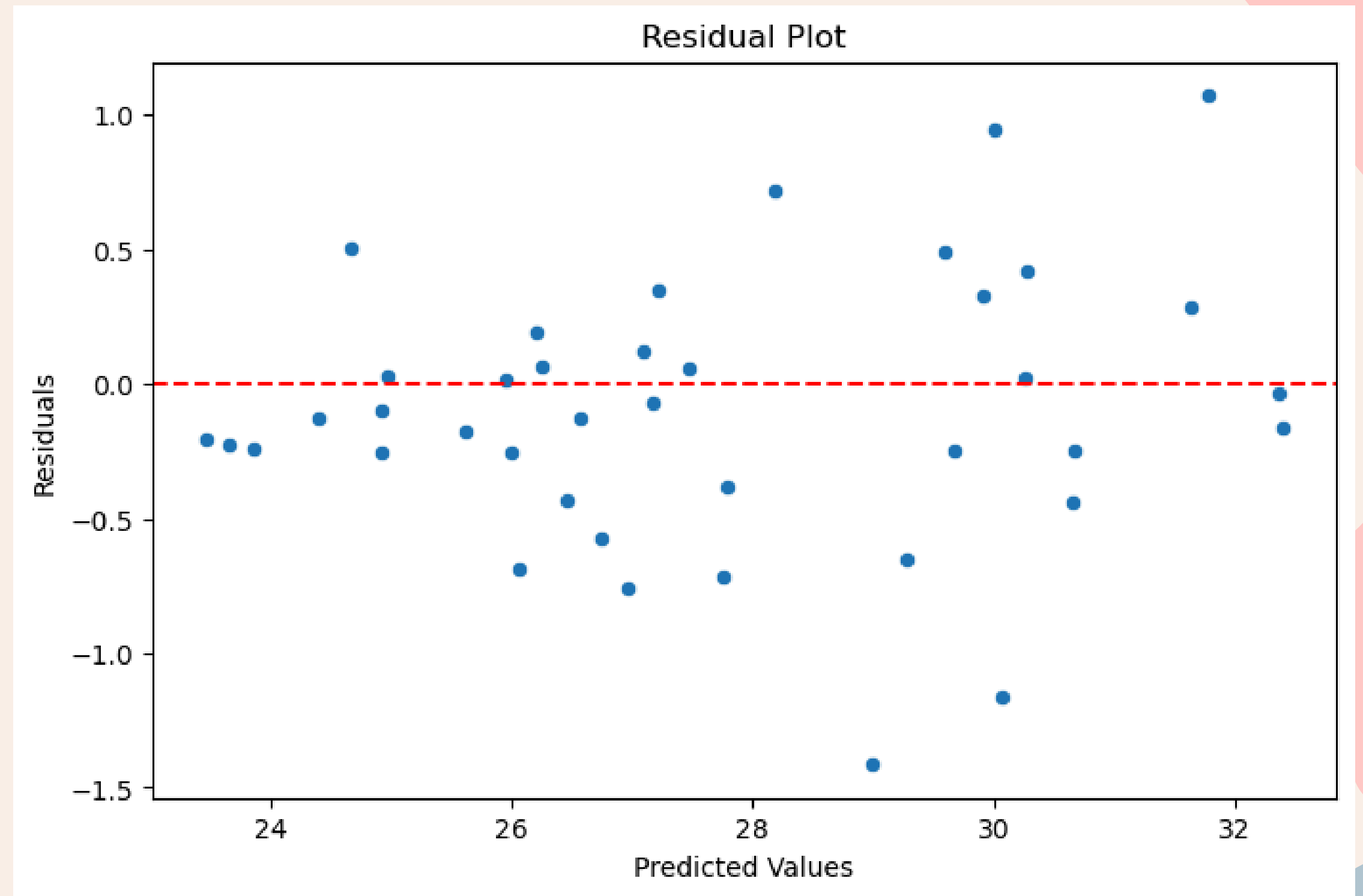
DATA TEST : PREDICT & HYPERPARAMETER TUNING

Best Parameter :
{'model__bootstrap': False,
'model__max_depth': None,
'model__max_features': 'sqrt',
'model__min_samples_leaf': 1,
'model__min_samples_split': 2,
'model__n_estimators': 200}

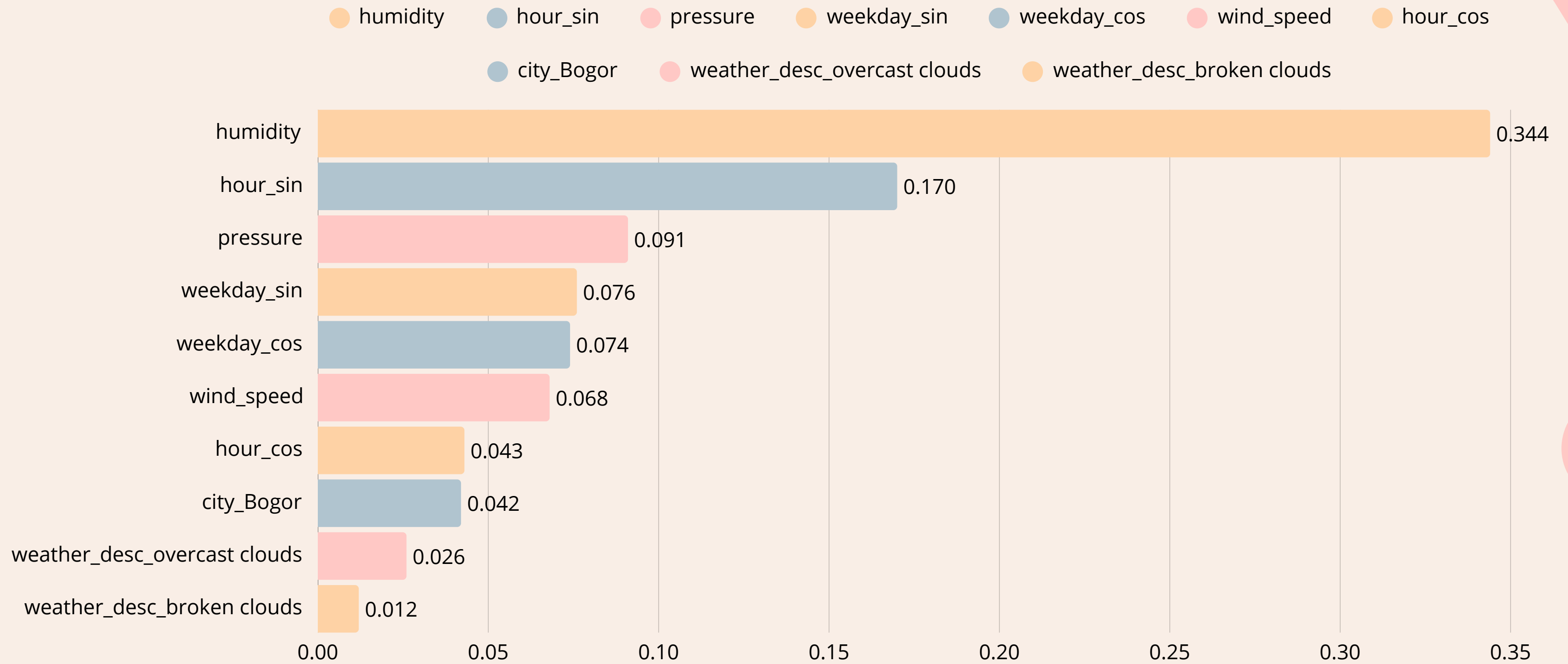


RESIDUAL PLOT

- Pola residual acak :
Titik residual tersebar acak di sekitar garis nol → model tidak melewati pola penting.
- Variasi stabil :
Penyebaran residual konsisten dan tidak ada indikasi heteroskedastisitas → model stabil.
- Outlier dibiarkan :
Beberapa outlier terdeteksi, tapi jumlahnya sedikit dan tidak berdampak signifikan terhadap model (Random Forest tahan terhadap outlier).



FEATURE IMPORTANCE



CONCLUSION

- **Akurasi Tinggi :**

Model Random Forest memberikan hasil sangat baik (RMSE: 0.5071, MAE: 0.3820, R^2 : 0.9623).

- **Stabil & Andal :**

Residual tersebar acak, tidak ada pola khusus, tidak ada indikasi heteroskedastisitas.

- **Tahan Outlier :**

Outlier sedikit & mencerminkan kondisi nyata – tidak mengganggu performa.

- **Fitur Penting :**

Kelembaban paling berpengaruh, diikuti waktu & tekanan udara. Fitur kategori (kota, cuaca) berpengaruh kecil.

RECOMMENDATION

- **Pengambilan Keputusan Lokasi**

Gunakan prediksi suhu harian untuk saran indoor/outdoor + perlengkapan pendukung (tenda, AC).

- **Optimasi Perlengkapan Acara**

Prediksi suhu bantu menentukan kebutuhan kipas, pemanas, air cooler, dll.

- **Integrasi ke Sistem Internal**

Tambahkan ke dashboard atau sistem konsultasi untuk akses cepat & otomatisasi prediksi.

- **Monitoring & Retraining Model**

Evaluasi prediksi vs realita secara berkala, lakukan retraining tiap 6–12 bulan.

- **Eksplorasi Fitur Tambahan**

Tambahkan data seperti UV index, polusi, atau hari libur jika relevan dan tersedia.

The background features a light beige base with several abstract, organic shapes in orange, pink, and blue. The orange shapes are located in the top-left and bottom-left corners, with the top-left one containing white dots. The pink shapes are in the top-right and middle-right areas, with the top-right one featuring white concentric lines. The blue shapes are in the bottom-left and bottom-right corners, with the bottom-right one containing white dashed lines.

THANK YOU