Department of Information Technology

Vasyl Stefanyk Precarpathian National University

Ivano-Frankivsk, Ukraine

Methods and algorithms for forecasting the effective use of energy resources for household consumers

Roman Dyndyn, Ihor Lazarovych, Serhii Ishcheriakov

Methodology

Feature Engineering

Algorithms comparison

Raw dataset

1. Consumption
2. Date
3. City
4. Area
5. Residents

Feature Engineering

1. Normalized consumption per day.
2. Annual consumption trend.
3. The deviation of the mean (logarithm) consumption to the mean consumption of all neighbors.
4. Correlation between consumption and temperature.

Feature Engineering

5. Lagged consumption (1 month)
6. Lagged consumption (2 month)
7. Month Sine:

$$MonthSin = \sin\left(\frac{2*\pi*m}{12}\right),$$ (1)

where m - month index.

8. Month Cosine:

$$MonthCos = \cos\left(\frac{2*\pi*m}{12}\right),$$ (2)

where m - month index.

Feature Engineering

Area	Person	Lag1	Lag2	Correlation	MeanDailyConsumption	DailyConsumptionDeviation	Trend	Month_Sin	Month_Cos	Year_Sin	Year_Cos
44.0	2	109	107	-0.511820	3.817982	0.101135	-0.155979	5.000000e-01	-8.660254e-01	-6.447061e-13	1.0
44.0	2	107	121	-0.511820	3.817982	0.101135	-0.155979	8.660254e-01	-5.000000e-01	-6.447061e-13	1.0
44.0	2	121	121	-0.511820	3.817982	0.101135	-0.155979	1.000000e+00	6.123234e-17	-6.447061e-13	1.0
44.0	2	121	123	-0.511820	3.817982	0.101135	-0.155979	8.660254e-01	5.000000e-01	-6.447061e-13	1.0
44.0	2	123	104	-0.511820	3.817982	0.101135	-0.155979	5.000000e-01	8.660254e-01	-6.447061e-13	1.0
...	...	...	...	...	...	...	...	...	...	...	...
60.9	2	107	116	-0.564552	4.336590	0.236746	-0.089648	-1.000000e+00	-1.836970e-16	-1.305266e-12	1.0
60.9	2	116	132	-0.564552	4.336590	0.236746	-0.089648	-8.660254e-01	-5.000000e-01	-1.305266e-12	1.0
60.9	2	132	138	-0.564552	4.336590	0.236746	-0.089648	-5.000000e-01	-8.660254e-01	-1.305266e-12	1.0
60.9	2	138	138	-0.564552	4.336590	0.236746	-0.089648	1.224647e-16	-1.000000e+00	-1.468363e-13	1.0
60.9	2	138	138	-0.564552	4.336590	0.236746	-0.089648	5.000000e-01	-8.660254e-01	-1.468363e-13	1.0

Figure 1 – List of all the features used for forecasting

Algorithms comparison

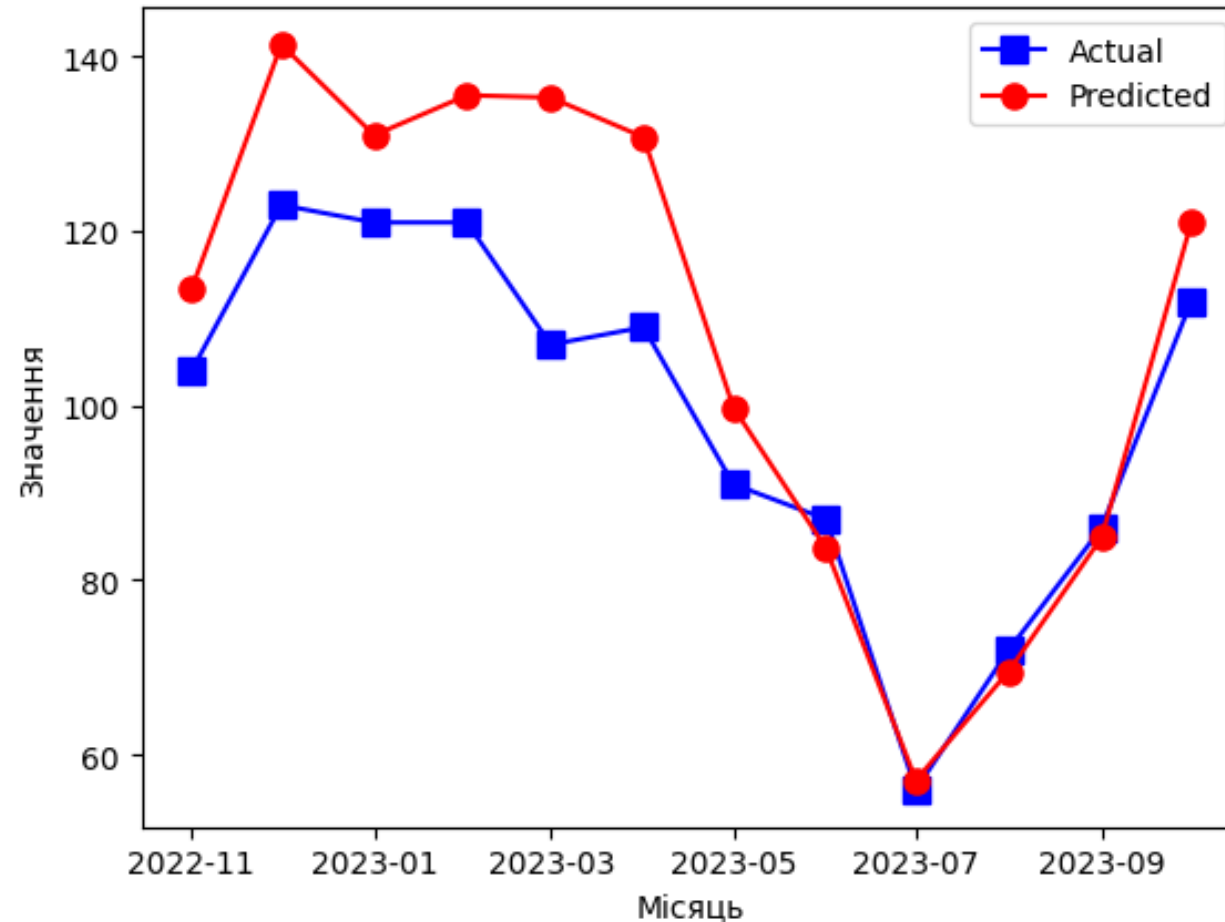| Method	MAE	MAPE
Linear regression	18.154	18.43%
Decision tree	22.333	22.25%
Random forest	17.137	17.13%
XGBoost	20.763	20.43%
Support vector machine	19.651	24.72%
Feedforward neural network	22.734	23.19%

One month ahead forecasting

Figure 2– Graph of predicted and actual electricity consumption for one user.

Results

As a result of the use of Random Forest algorithm and the application of the proposed approaches for feature engineering, it was possible to achieve an improvement in accuracy by 1% compared to other algorithms.

Future Research

Considering the small amount of data sample for model training, it is necessary to investigate the behavior of the proposed approach on more voluminous data. It is also advisable to investigate in more detail the influence of various holiday periods, such as Easter, on the consumption of energy resources.

Thank you for your attention