E-Adivino: A Novel Framework for Electricity Consumption Prediction based on Historical Trends

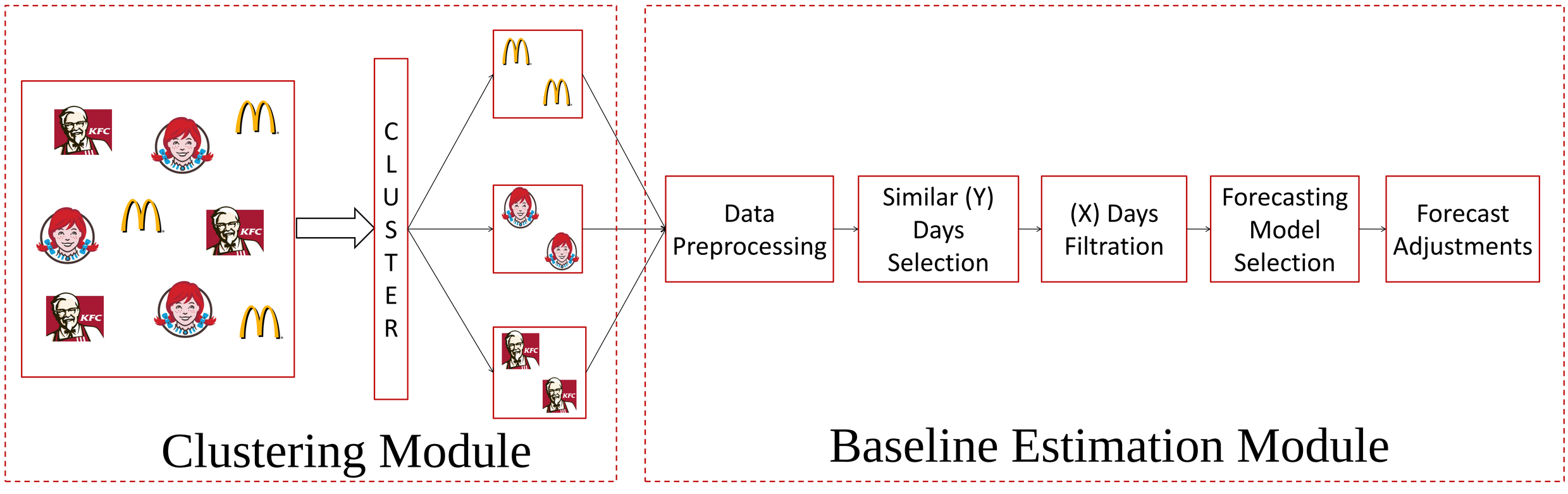
Shubham Saini†, Pandarasamy Arjunan†, Amarjeet Singh†, Ullas Nambiar‡
†IIIT Delhi
‡ EMC Corporation

Drawbacks of Current Baseline Estimation Models

❖Same baseline estimation technique is used by a DR program for all their customers.[1]
❖Wherever modeling is performed, different parameters are learned for each customer separately.[1]

Cluster similar loads and find the best baseline estimation model for each cluster

Simple Intuition: daily energy usage patterns for facilities with similar size, business hours, electrical equipment and geographical location are likely to be similar.



Test Data

❖IIIT Delhi Academic Campus
❖Over 200 smart meters across 7 buildings collecting data every 30 seconds
❖More than 2.5 billion data points collected till date

Accuracy Measure

Mean Absolute Scaled Error[3]

$$MASE = \frac{\sum_{t=1}^{T} |y_t - \hat{y}_t|}{TQ}, \; Q = \frac{\sum_{t=m+1}^{T} |y_t - y_{t-m}|}{T-m}$$

Where:
y_t: observed value
$\hat{y}_t$: predicted value
Q: stable measure of scale
m: period
T: length of time series

Results

Best Fit	0.41
Temporal Features Clustering	0.88
DTW Clustering	0.95
NYISO Baseline	1.46

Parameter	Values
Clustering Algorithm	Partitioning Around Medoids
Similarity Measure for Clustering	Temporal Features[2] Dynamic Time Warping
Forecast Horizon	{8,24} Hours
Similar (Y) Days Selection Criteria	Previous Business Days Previous Same Days
Number of (Y) Similar Days	{4,7,10} Days
Filtering (X) Days Criteria	{High X, Middle X}
X:Y Ratio	{0.6, 0.8}
Forecasting Model	Hybrid ARIMA+ANN Simple Averaging Adjusted Averaging

Future Work

❖Efficient scaling of loads clustering module to account for real life scenario, where a utility has millions of customers.
❖Making use of additional context information to improve the prediction accuracy.

References

[1] Clifford Grimm and DTE Energy. Evaluating baselines for demand response programs.
[2] Xiaozhe Wang, Kate Smith, and Rob Hyndman. Characteristic-based clustering for time series data.
[3] Rob J Hyndman and Anne B Koehler. Another look at measures of forecast accuracy.