Lessons and Insights from Super-Resolution of Energy Data

Rithwik Kukunuri
UG Student, IIT Gandhinagar
kukunuri.sai@iitgn.ac.in

Nipun Batra
IIT Gandhinagar
nipun.batra@iitgn.ac.in

Hongning Wang
University of Virginia
hw5x@virgnia.edu

Motivation

Studies have shown that consumers of electricity can save up 15% of their bills when provided with a detailed appliance wise feedback [1]. Energy super-resolution refers to estimating energy usage at a higher-sampling rate from the lower sampling rate. We mainly focus on predicting the hourly reading of a home, using the daily usage (which can be noted down by the users from the meter). This predicted usage can be used by the consumers to identify the times of the day, which are contributing more to electricity usage and help them optimize their usage. This is analogous to image super-resolution, where the zooming out factor equals 24.

Problem definition

Throughout the paper we will be using the following notation: H - Number of homes; D - Number of days; $X \in R^{H \times D}$ - Denotes low resolution matrix (Aggregate); $Y \in R^{H \times D \times 24}$ - Denotes high resolution matrix; $P \in R^{H \times D \times 24}$ - Denotes weights matrix; Weights matrix is same as the matrix which stores the proportion of electricity consumed on a particular day. For the h^{th} home and the d^{th} day, the matrix $\forall_i P_{h,d,i} = \frac{Y_{h,d,i}}{X_{h,d}}$

Approach - Triplet learning

Let $L(i,j)$ denote $X[i, j-K : j+K]$, which is a vector of length $2K$+1. It stores the K past and K future neighbor aggregate readings in a home i during day j. We can refer to this a neighborhood vector for the ith home for jth day. An embedding network takes $2K$+1 dimension vector as input and outputs an vector of dimension N. The embedding network can be configured with various options such as normalization of output and positive activation of output.Consider (i,x),(j,y),(k,z), where each tuple denotes a home and day pairs. Let $V(i,x)$ denote the embedding vector generated using $L(i,x)$. We define similarity functions which are specified in $Equation(1)$

$$
\begin{aligned}
a &= \left\|P[i,x] - P[j,y]\right\|_2 \\
b &= \left\|P[i,x] - P[k,z]\right\|_2 \\
c &= \left\|V[i,x] - V[j,y]\right\|_2 \\
d &= \left\|V[i,x] - V[k,z]\right\|_2
\end{aligned}
\tag{1}
$$

The functions in $Equation(1)$denote the similarity of the given tuples in the super-resolution usage. The losses in Table 1 ensure that tuples that are similar in the weights space are also similar in the embedding space. After the embedding network finished training, we generate the embeddings for each of the test samples. Then we find k nearest training samples using the embeddings and use the weights of the closest samples as the weights for the test sample. The k and margin are chosen using a valiation set.

Baselines

Uniform weights: In this model, we assume that the usage of a home is constant throughout the day.

Table 1: Triplet loss on different inputs and embeddings

Condition 1	Condition 2	Loss
a>b	c>d+margin	0
a>b	c<d+margin	$(d + margin - c)^2$
a<b	c+margin<d	0
a<b	c+margin>d	$(c + margin - d)^2$

Table 2: MAE on different models

Model	MAE
Uniform weights	421
Mean weights	338
K-Neighbours weights	357
Tensor Decomposition	356
Triplet Learning	348

Mean weights: In this model, we find the 24-dimension weights vector that minimizes the super-resolution error across all homes for all days. This is the mean of 24 dimension weights vector across all homes and days computed from the training vector.

K**-nearest neighbors weights estimation**: Estimates the weights for a home by finding the nearest neighbors using the aggregate vector. The weights will be multiplied the aggregate usage of the house to estimate super-resolution usage.

Tensor Decomposition: We create a tensor T of shape (H, D, 25) where one channel corresponds to aggregate and remaining twenty-four dimensions belong to super resolution usage. There are estimated using Low Rank Parafac Tensor Decomposition into r components. $\hat{T} = (Hxr) \otimes (Dxr) \otimes (25xr)$ where $\otimes$ denotes the outer product operator.

Results

We chose homes from Dataport [3] dataset from June. There are a total of 51 training homes, 7 Validation homes, and 13 testing homes. Mean absolute error is the metric chosen to evaluate the performance of a model.

Analysis

The reason for similar results for all models (except uniform weights) can be because of the variance in the electricity usage of a home at hourly level. The high variance is causing all models to cause similar errors across homes and averaging out the error, thereby bringing it closer to mean.

Future work

In future we would like to explore the area of super-resolution plus disaggregation, where we also estimate the appliance usage at an hourly level. We plan to use Generative adversarial networks [2] which are state-of-the-art methods in the image super-resolution.

REFERENCES

[1] K. Carrie Armel et al. "Is disaggregation the holy grail of energy efficiency? The case of electricity". In: *Energy Policy* 52 (2013). Special Section: Transition Pathways to a Low Carbon Economy, pp. 213–234. ISSN: 0301-4215. DOI: https://doi.org/10.1016/j.enpol.2012.08.062. URL: http://www.sciencedirect.com/science/article/pii/S0301421512007446.

[2] Christian Ledig et al. "Photo-realistic single image super-resolution using a generative adversarial network". In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017, pp. 4681–4690.

[3] Oliver Parson et al. "Dataport and NILMTK: A building data set designed for non-intrusive load monitoring". In: *2015 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*. IEEE. 2015, pp. 210–214.