EMERGING TECH CONFERENCE – Edge Intelligence

Volume 01, 2022, Page 154 – 160

**Proceedings of Emerging Tech Conference:
Edge Intelligence 2022**

Energy Disaggregation on Real Time Ground Truth Data

Athanasios Argyropoulos[a], Aristeidis Karras[a], Christos Karras[a], Stelios Koutroubinas[b]*, Kostas Tsichlas[a], Spyros Sioutas[a]

[a]*Departerment of Computer Engineering & Informatics, University of Patras, 26504 Patras, Greece*
[b]*Meazon S.A., Ippodamou Str., 26442, 26504 Patras, Greece*

s.koutroubinas@meazon.com

Abstract

Deep neural networks are regarded as the fundamental axis of the artificial intelligence research in the present day. Recent years have seen an increase in the volume of information, which has led to the creation of sophisticated deep learning models. These models are able to draw inferences from massive data sets that aid in understanding and addressing issues. In contrast, significant progress in hardware and electronics has led to the development of inexpensive IoT data gathering devices for everyday usage. One such intelligent gadget is the electric meter, which is installed in a home's central panel and records electrical data. A contemporary challenge in electricity consumption is the separation of total electrical energy into states that indicate the use of particular home appliances. With the use of deep neural network models and the data obtained by smart meters, it is possible to monitor the power usage of each individual device, resulting in energy savings and a greater concern for the environment. Consequently, the objective of this study is to construct regression-based deep learning models that segregate electricity from actual data gathered by smart meters of the business MEAZON S.A.

1. Introduction

Nowadays, power conservation is a topic of considerable public concern. It is extremely useful for saving money and being environmentally friendly. The advancement of technology has resulted in the creation of smart data gathering devices for common users, which facilitate the construction of robust models of consumption, such as smart energy meters.

The vast amount of information in the contemporary period has led to the development of powerful data analysis algorithms and engineering and deep learning approaches for inference. This is how the demand for energy a- analysis arose, which examines energy data with an intelligent approach and draws findings that increase the effectiveness of electric energy management in homes and businesses. The difficulty of the thesis in energy analysis is the separation of energy (energy disaggregation) or non-intrusive monitoring of load. The concept of separation entails the measurement of just the household's voltage and current for the goal of isolating the consumption current of the distinct states of the devices [1].

The power consumption of the home is measured in watts, which is often used for measuring device use (W). The purpose is to forecast the electrical demand of the remaining vessels in a time series fashion by using smart meters installed in the house's central panel to measure the electrical load and for energy disaggregation [2][3] purposes. Notably, the solution to the issue does not include putting a separate meter on each device due to the significant cost and complexity involved [4].

In conclusion, the detailed information of the instantaneous electrical load of individual apartments, which is provided by smart meters and technicians utilizing machine learning, enables the estimation of the overall consumption profile of a new flock, thereby allowing electricity to be used for a variety of policy purposes.

1.1 Objective

The objective of this work is to separate the overall energy signal from the energy signals of specific home appliances. In order to estimate the electrical load of a certain electrical device, we use deep learning approaches using combinatorial neural network topologies. We develop several types of autoencoders, including denoising autoencoders and altering autodecoders. Python programming was used to create the system utilising the Tensorflow and Keras frameworks. As gadgets - goals were established:

- Oven electric
- Air conditioning
- Water heater

The data utilised is actual, gathered from two distinct residences over an extended period of time, and the measurements were performed by real measuring devices of time (Smart metres), which were put in the residence's central panel and provided by Meazon S.A. Experiments conducted in the learning model used the entire amount of power used by the metre (Active Power(W)) as input and reconstructed the energy consumption signal of a particular device. Finally, the findings and conclusions of the experiments are provided.

2. Preliminaries

Whether we are aware of it or not, machine learning is an integral part of our daily life since we use it daily via our gadgets. What, though, is the fundamental concept behind the Machine Learning idea?

Machine learning is a subject of computer science that evolved from the study of pattern recognition and the theory of computational learning to artificial intelligence [1]. Using machine learning algorithms, complicated data patterns are discovered automatically. Using a problem's latent patterns and awareness of them, future occurrences may be predicted.

Traditionally, software engineers created solutions to a problem by combining data with human-made rules. Instead, engineering learning use data and reactions to identify the rules underlying an issue. The three major types of machine learning are supervised, unsupervised, semi-supervised, and reinforcement learning.

2.1. Neural Networks

A layer of nodes comprises an input layer, one or more hidden layers, and an output layer in artificial neural networks. Each node, or artificial neuron, is interconnected and has a weight and threshold linked with it. If the output of a node exceeds the threshold value, that node becomes active and transfers data to the next tier. If it is less than the value below-layer, no data is sent to the next network layer. They rely on data training to learn and increase their precision over time.

After the training procedure is complete, an artificial neural network enables sorting and grouping. Compared to manual re- knowledge from human professionals, voice recognition and picture recognition tasks may be completed in a matter of minutes.

Each node in the artificial neural network has its own linear regression model, comprised of input data, weights, bias (or threshold), and output. Simply a level of admission is chosen, and weights are provided. These weights assist in determining the significance of each variable, with greater weights contributing more heavily to output than smaller ones. All inputs are multiplied by their respective weights before being added together. The output is then processed by an activation function that determines the output. If its output exceeds a predetermined threshold, the node is "triggered" (or activated) by transferring the data to the subsequent network layer. This causes a node to leave at the entry of the subsequent node. In the diagram below, we can see how the network nodes are connected. The transmission of data from one layer to the next layer characterizes this neural network as a feedforward network. Hence, we have:

$$\sum_{i=1}^{m} w_i x_i + \text{bias} = w_1 x_1 + w_2 x_2 + w_3 x_3 + \cdots + w_m x_m + \text{bias} \quad (1)$$

3. Data Preprocessing

Data preparation consists of the processes required to convert or encode the data so that it may be readily analyzed by a computer. It may refer to the process of manipulating or eliminating data before to their usage to guarantee or enhance performance. Frequently, the data that need pre-processing are out of their pricing range, such as (for example, Age: -10), impossible data combinations (for example, Gender: Male, Pregnant: Yes), and missing values. Before undertaking any analysis, the representation and quality of the data must thus be considered first and foremost. [9] [10]. It is not feasible to train algorithms with raw data. Therefore, we preprocess the datasets derived from MEAZON S.A. There are four primary improvements that we perform on our dataset among:

- Filling in missing values
- Identifying outliers
- Smoothing of noisy data
- Fix inconsistent data

As per each technique, certain actions take place. Initially for the missing values the mean of each feature is used along with binning methods. Hence, the data is more equally distributed than the raw data we had previously. Moreover, for the outlier detection we perform linear regression in order to see whether and which data fall far from the regression line. These data are desegregated from the training phase of the model. Additionally, to reduce or eliminate noisy data we use smoothing algorithms. Lastly, we normalize values which helps us to represent data in a shorter range so as to have a more realistic view of the features of the dataset.

4. Methodology

4.1. De-noising Autoencoder

In this section we introduce the two techniques followed in order to construct a De-noising Autoencoder and a Variational Autoencoder. Autodecoder design permits the representation of input data as equivalent output data. In this scenario, the model may be used to denoise the data. The objective is to distort the information judged to be noise while keeping the relevant information as intact as feasible. Such models capable of removing noise from autoencoders is named denoising autoencoders (DAE). The topic is extended as a problem of supervised learning with damaged input data and noise-free goal data. The objective is to recreate the data using the neural network as a noise filter. During the encoding process, bottlenose extracts the sensory properties that make up the data, and the encoder reconstructs the data. The model is tasked with minimizing the difference between the modified input and the original input.

4.2. Variational Autoencoder

The variable autoencoder (VAE) has the capacity to utilise latent space values in a probabilistic way, since it defines the probability distribution of space in order to make a prediction. Samples the latent space that forms a distribution for each distinct class. The model consists of an encoder and a decoder, however for its probabilistic dimension, the input is encoded as a distribution over latent space rather than as a point. Then, when decoding from the latent space, a random sample from each latency distribution is taken to form one vector for the decoder's input.

For every sample from the latent distributions, the entry is precisely reconstructed by the decoding model poet. Consequently, elements that are near to one another in the latent space should correlate to quite comparable reconstructions. The coder is defined probabilistically as p (z|x), which is the distribution of the code-of copied variable given the decoder. The decoder is configured probabilistically as p(x|z), where defines the decoded variable's distribution given the encoding scheme. The encoded representations of z in the latent space adhere to the previous distribution p(z). Based on Vapses' theorem, p(z|x) is represented as follows:

$$p(z \mid x) = \frac{p(x \mid z)p(z)}{p(x)} \tag{2}$$

Where p(x) can be expressed as:

$$p(z \mid x) = \frac{p(x \mid z)p(z)}{p(x)} \tag{3}$$

In the figure below we see the probabilistic encoder and decoder, as well and the latent space.

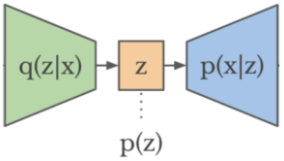

Figure 1: Probabilistic Encoder/Decoder

This technique is computationally extremely "expensive," hence an estimate of p(z|x) is derived from a known distribution q(z|x) (e.g., Gaussian) in order to approximate the extraction results. As

stated before, this is accomplished by minimizing the Kullback-Leibler divergence, which approximates the distance between two distributions as below.

$$\min K\,L\big(q(\,z\mid x\,)|p(\,z\mid x\,)\big) \tag{4}$$

The loss function of the model consists of two terms, one that minimizes the reconstruction error and a second term that maximizes the similarity of the distribution q(z|x) real distribution p(z), which assumes that it follows a unitary Gaussian distribution for each dimension j as below.

$$\mathcal{L}(x,\hat{x}) + \sum_{j} K\,L\left(q_j(\,z\mid x\,)|p(z)\right) \tag{5}$$

Since we assume that the distribution follows a Gaussian distribution, the mean and variance of the distributions from the latent space are described by two vectors. If a real Gaussian model were to be developed for a mixture, a covariance matrix representing the mode associated with each of the dimensions would need to be generated. The sampling method is crucial since the model is required to execute backpropagation and must be trained. At this stage, the reparameterization approach is used, where he advises selecting a random sample e from a unit Gaussian distribution and then shifting the sample by the mean m and variance s of the latent distribution, respectively.

$$z = \mu + \epsilon \cdot \sigma \tag{6}$$

Therefore, reparameterization improves the parameters of the error-free distribution stages during the sample procedure. Where will we have:

$$z \sim q_{\mu,\sigma}(z) = \mathcal{N}(\mu,\sigma) \tag{7}$$

And:

$$\epsilon \sim \mathcal{N}(0,1) \tag{8}$$

5. Results

The data used in the experiments are derived via the installation of the DinRail 3-phase WiFi plug which was attached on the central panel within a house and collected electricity information from each user's profile. The sensor measured the characteristics and total energy signal, but and two other sensors were installed to measure the electric furnace in one home and the water heater. In the second household, data were collected from the electric oven and the heating system furnace and the air conditioner. The appliances were measured separately to collect ground truth data. The measurement period for the first household was from 01-06-2021 to 01-10-2021 and for the second one 07-05-2021 to 13-05-2021.

The data were preprocessed prior to analytics and more specifically, the negative values that the meter receives. Moreover, some readings of the smart meter, which ranged from 0 to 10 watts, were set to zero since they were latent. As it can be observed in the images, all data were standardized to a value range of 0 to 1. On a scale from 0 to 1, the following figures depict the overall electrical signal of the target device. The first batch of data has the following wattage time series structure. The goal variable was then assigned to the electric heater, followed by the air conditioner. The following figure depicts the electrical signal of the target device (electric oven) and the overall electric current.

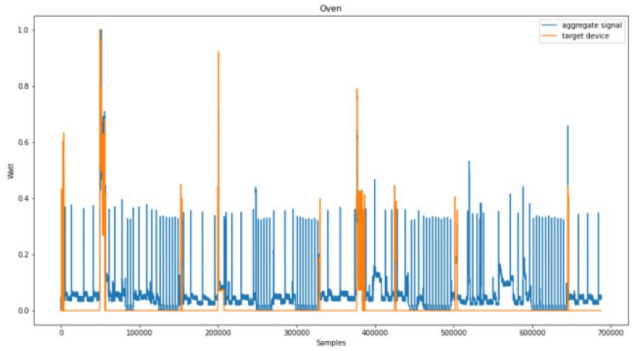

Figure 2: Electric oven and total current signal

The important de-noising autoencoder architecture from the paper Neural NILM: Deep Neural Networks Applied to Energy Disaggregation was the first deep learning model deployed [5]. It comprises of a one-dimensional input convolutional layer and three dense layers for encoding and decoding initiation in the hidden layer. Then, in the decoding, there is another convolutional layer, followed by a Flatten layer, and a node for the output of the regression in the output layer.

After a huge number of experiments and evaluations, the most effective pro-parameters for analysis were chosen. Linear activation functions were utilized for convolutional layers while ReLu activation functions were employed for dense levels. The output layer used a linear activation function. Adam, with a learning rate of 0.0001, was the most effective algorithm for optimization after a significant number of tests. The mean absolute error was selected as the loss function.

Electrical Appliance	MAE	MSE	RMSE	Loss
Oven	0.0106	0.0017	0.0417	0.1672
Air Condition	0.0554	0.0279	0.1618	0.027

Table 1: Results of first DAE data set

Electrical Appliance	MAE	MSE	RMSE
Oven	0.0099	0.0011	0.0344
Air Condition	0.0503	0.0261	0.1618

Table 2: Results of first **CVAE** data set

6. Conclusions

Non-intrusive load monitoring is an intriguing and contemporary problem that now preoccupies a significant number of scientists. The energy disaggregation results in energy savings, the avoidance of electrical grid overload, and the creation of a smart power grid. In this work, e-energy was separated using the regression approach so that a deep learning model can distinguish in real time the overall from the individual device consumption and identify it with the highest possible accuracy for each user profile. The data utilized was acquired from MEAZON S.A. meters, which

were installed in the central panel of each home. In the first, data were collected from the electric oven and air conditioner. ko, and in the second, the electric oven and water heater. As previously noted, the intention was not to deploy a separate sensor on each target device, but rather to forecast the consumption of the entire electrical signal relative to the signal of each individual device.

References

[1] Hart, G. W. (1992). Nonintrusive appliance load monitoring. Proceedings of the IEEE, 80(12), 1870-1891.

[2] Lykothanasi, K.K., Sioutas, S., Tsichlas, K. (2022). Efficient Large-Scale Machine Learning Techniques for Rapid Motif Discovery in Energy Data Streams. In: Artificial Intelligence Applications and Innovations. AIAI 2022. IFIP Advances in Information and Communication Technology, vol 646. Springer, Cham. https://doi.org/10.1007/978-3-031-08333-4_27

[3] Konstantopoulos, C., Sioutas, S., Tsichlas, K. (2022). Machine Learning Techniques for Regression in Energy Disaggregation. In: Artificial Intelligence Applications and Innovations. AIAI 2022. IFIP Advances in Information and Communication Technology, vol 646. Springer, Cham. https://doi.org/10.1007/978-3-031-08333-4_29

[4] Koutroumpina, C., Sioutas, S., Koutroubinas, S., & Tsichlas, K. (2021). Evaluation of features generated by a high-end low-cost electrical smart meter. Algorithms, 14(11), 311

[5] Kelly, J., & Knottenbelt, W. (2015, November). Neural nilm: Deep neural networks applied to energy disaggregation. In Proceedings of the 2nd ACM international conference on embedded systems for energy-efficient built environments (pp. 55-64).