Yousra Gaimes

Deep Learning Applications in Non-Intrusive Load Monitoring

Supervised by:

DR. ANTÓNIO EDUARDO DE BARROS RUANO

UNIVERSIDADE DO ALGARVE
Faculdade de Ciencias e Tecnologia

2021

Work Authorship Declaration

I declare to be the author of this work, which is unique and unprecedented. Authors and works consulted are properly cited in the text and are in the listing of references included.

Yousra Gaimes

……………………………

Acknowledgements

To my parents, my brothers, my family, Hélder Gonçalves and all those that one way or another contributed to my personal, social, and professional evolution.

A Very special thanks to my Professor Antonio Ruano and my colleague Habou Laouali Inoussa.

Abstract

Within the frame of the project Non-Intrusive Load Monitoring for Intelligent Home Energy Management Systems, this work will present a deep learning application in non-intrusive load monitoring on a case study in a residential home in in Gambelas, Faro in the Algarve region south of Portugal. This work has for a goal to detect type 2 appliances in different houses. For the sake of this study, two models will be trained:

- Convolutional Neural Network
- Long Short-term Memory Recurrent Neural Network

on three datasets:

- UKDale
- REDD
- Data from the Portuguese private residential house from the project NILM for IHEMS.

Keywords:

NILM, CNN, LSTM, UKDale, REDD, NILM for IHEMS.

Resumo

No âmbito do projeto Monitorização de Carga Não Intrusiva para Sistemas Inteligentes de Gestão de Energia Doméstica, este trabalho apresentará uma aplicação de aprendizagem profunda na monitorização de carga não intrusiva num estudo de caso numa casa residencial em Gambelas, Faro na região sul do Algarve de Portugal. Este trabalho tem por objetivo detectar eletrodomésticos tipo 2 em diferentes residências. Para fins deste estudo, dois modelos serão treinados:

- Rede Neural Convolucional
- Rede Neural Recorrente de Memória Longa de Curto Prazo

em três conjuntos de dados:

- UKDale
- REDD
- Dados da habitação privada portuguesa do projecto NILM para IHEMS.

Palavras-chave:

NILM, CNN, LSTM, UKDale, REDD, NILM para IHEMS

Table of Content

Table of Figures & Tables

1 Chapter 1: Introduction

1.1 Non-Intrusive Load Monitoring (NILM)

As the cost of energy is increasing and climate is changing, awareness of energy consumption is rising among households and industrials. Our lifestyle is dependent on a large quantity of energy consumption. Almost all our daily tasks are using electric appliances. This pushed the humanity to look for ways to understand the user consumption and monitor it. It is very important for both the consumer and the provider to understand the patterns of consumption so that they can take actions and try to optimize it. For example, if users understand their home appliances consumption, they can draw conclusions relatively to the appliance consumption rate, when to efficiently operate some appliances and how often. Awareness on appliances consumption will help the users to get a more concrete sense of what they are consuming through a proper monitoring solution, such as load monitoring. This monitoring technique was first applied in the industrial buildings then lately to households.

Load Monitoring consists of two types “Intrusive Load Monitoring” and “Non-Intrusive Monitoring”. The intrusive one consists of installing smart meters to every plug of each appliance. The major drawbacks of this first type of load monitoring are that it intrudes the household and requires additional hardware that most of the time is expensive. The second type, which will be the topic of this thesis, is the Non-Intrusive Load Monitoring (NILM). In this later type, it is not required to get into the house and intrude people’s privacy; all what is needed are the readings from the general meter of the house.

Non-intrusive load monitoring was first introduced in 1992 by G. W. Hart (HART, G. W. 1992). It is a method that, based on household power data acquisition, disaggregates the total power into appliances’ single consumptions. This method is based on algorithms that obtain all the input data from one meter connected to the grid. This makes it a non-costly and non-intrusive method that proved its efficiency in multiple researches that will be discussed further on in this dissertation.

Energy monitoring using non-intrusive load monitoring allows multiple outcomes (Gopinatha et al., 2020):

- It can be used to detect and identify appliances.
- It can be used to give feedback to users on consumption levels of their appliances.

- It can identify appliances in houses without invading the house.
- It can give feedback on the appliances performance and degradation over ageing.
- It can analyze appliances behavior and detect anomalies.
- It can be used to control the operation of appliances and the inverter, if a PV installation is available.

Non-intrusive Load Monitoring is an approach that allows the energy consumer and the energy provider a cheap, efficient, and a simple access to very important data to have insight on the energy consumption.

We can also define the non-intrusive load monitoring, also known as power disaggregation, as an approach trying to solve to following equation:

$$P_H = \sum_{i=1}^{A} P_i + \varepsilon$$

P_H: Total power consumed in the house

P_i: Power consumed by appliance i

ε: Noise

A: Total number of appliances

To visualize the general concept of Non-intrusive load monitoring technique, figure 1 shows how the appliances are connected to the main power supply of the building, which is the main input of the NILM algorithms.

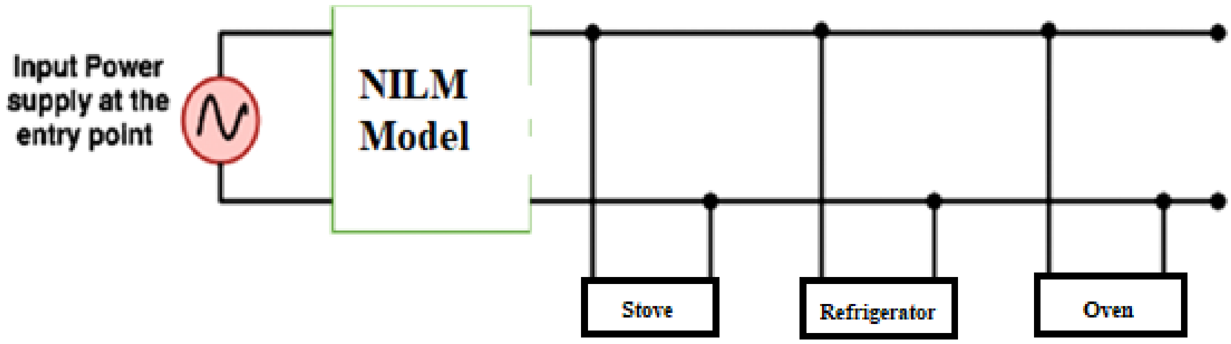

Figure 1 NILM general concept (Gopinatha et al., 2020)

To deep dive into what happens in the energy consumption, the following figure (figure 2) shows the power consumption of different appliances in a household.

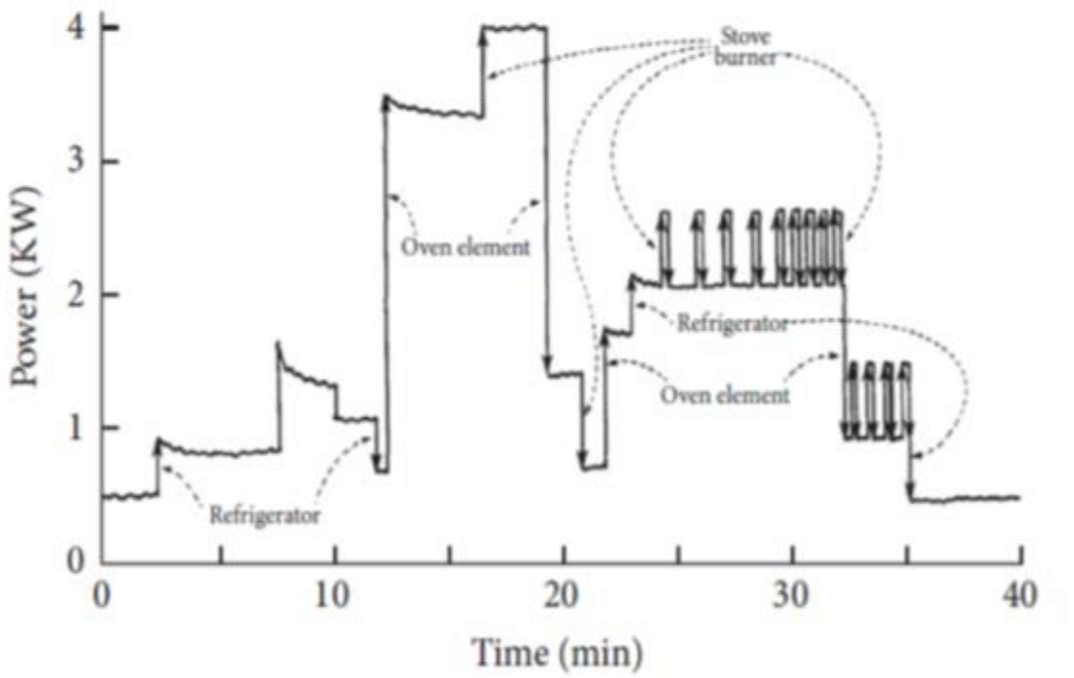

Figure 2 Total power consumption of a house showing different appliances (G. W. Hart, 1992)

1.1.1 Appliance Types

Generally, a wide variety of appliances is used in households. Thus creating a big challenge for NILM scientists, who opted for the following appliances' categorization (figure 3) to help them better handle the issue on hand.

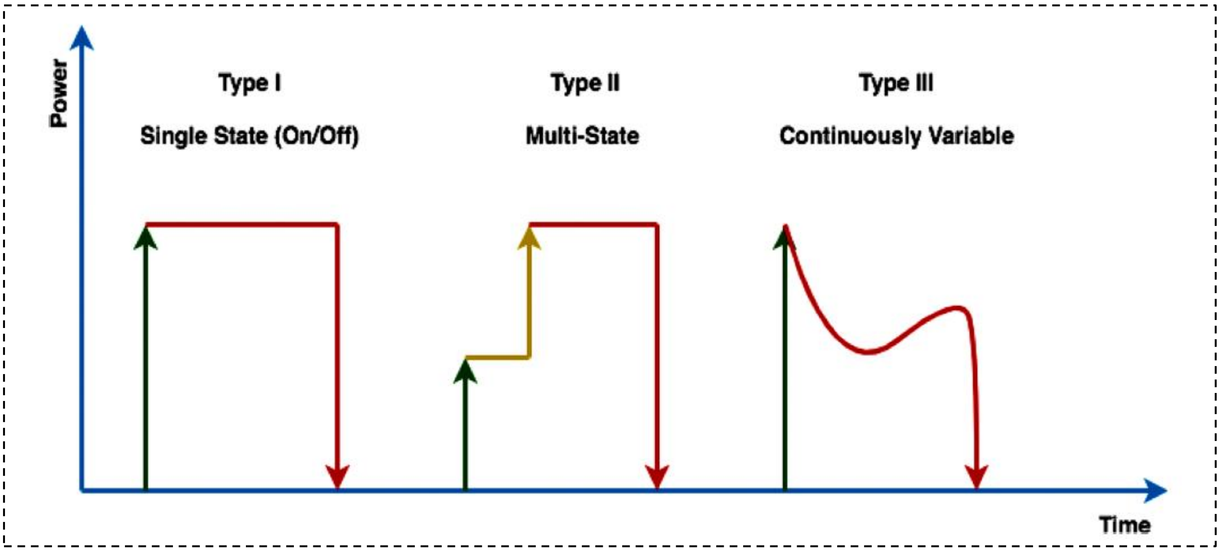

Figure 3 Appliances types (Gopinatha et al., 2020)

Type-I:

Appliances operating on a single state (ON/OFF).

Example: Lamps, toaster.

Type-II:

Appliances operating with finite-states. These have a finite number of operating states, and can be denoted as Finite State Machines (FSM). The transition between the states usually cannot be manually controlled.

Example: Washing machine, Stove burner, and hair dryer.

Type-III:

Appliances known as Continuously Variable Devices (CVD). They operate on variable power with no fixed number of states.

Example: Power drill, Dimmer lights

Type-IV:

Appliances cyclically on/off. These appliances have a periodic nature.

Example: House alarm, Electric heater.

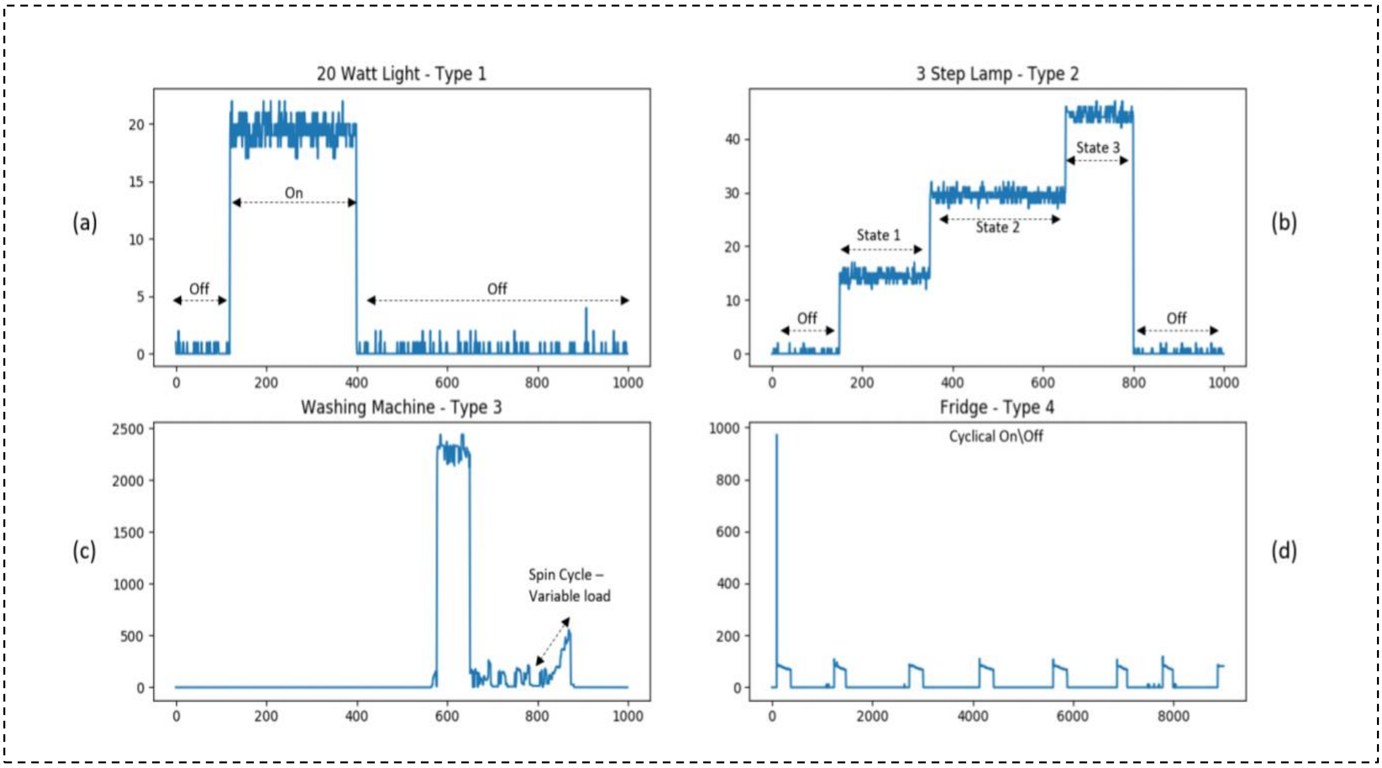

Figure 4 Power Signatures of different appliance types

(a) 20 Watt light (b) Lamp with 3-user controlled intensity settings

(c) Washing machine (d) Fridge

Understanding the types of appliances is crucial in solving NILM problems. Every type has a very distinct signature that defines it and allows the algorithms to learn to recognize it.

1.1.2 Power signal

Nowadays, the electrical power, almost in every household, comes from the grid as an Alternating Current (AC). This household current has an oscillating nature and can go negative at moments when power is returning to the grid. For this reason, AC is separated in two main types:

- Active Power (P)
- Reactive Power (Q)

1.1.2.1 Active Power

It is the Real power used in an AC Circuit. It is measured in Kilowatt (kW).

1.1.2.2 *Reactive Power*

It is the power that moves back and forth to the grid from the household. It is measured in Kilo volt-ampere reactive (kVAR).

In this research, we will be working with the Active Power only.

1.1.3 Datasets

In the Non-Intrusive Load Monitoring field many datasets are available for research and development. The data is collected from different locations around the world during various periods of time and different sampling rates. We can cite for example:

- **Reference Energy Disaggregation Data Set (REDD):** "*consists of whole-home and circuit specific electricity consumption for a number of real houses over several months' time. For each monitored house, we record the whole home electricity signal recorded at a high frequency (15kHz); up to 24 individual circuits in the home, each labeled with its category of appliance or appliances, recorded at 0.5 Hz; up to 20 plug-level monitors in the home, recorded at 1 Hz, with a focus on logging electronics devices where multiple devices are grouped to a single circuit*." (Kolter and Johnson, 2011)

- **UK Domestic Appliance-Level Electricity (UK-DALE):** "*an open-access dataset from the UK recording Domestic Appliance-Level Electricity at a sample rate of 16 kHz for the whole-house and at 1/6 Hz for individual appliances. This is the first open access UK dataset at this temporal resolution. We recorded from five houses, one of which was recorded for 655 days, the longest duration we are aware of for any energy dataset at this sample rate*." (Kelly and Knottenbelt, 2015b). Some details of this dataset are present in the following table:

House	*1*	*2*	*3*	*4*	*5*
Building Type	End of terrace	End of terrace	-	Mid-terrace	Flat
N° of Occupants	4	2		2	2
Total Number of Meters	54	20	5	6	26

Sample Rate of Mains Meters	6s	6s	6s	6s	6s
Date of 1st Sampling	09/11/2012	17/02/2013	27/02/2013	09/03/2013	29/06/2014
Date of Last Sampling	05/01/2015	10/10/2013	08/04/2013	01/10/2013	13/11/2014
Duration (Days)	786	235	39	206	137
N° Appliances	53	18	4	11	24
Average Main consumption per day (kWh)	7.64	7.17	-	-	13.75

- **Plug-Level Appliance Identification Dataset (PLAID):** "*includes current and voltage measurements sampled at 30 kHz from 11 different appliance types present in 56 households in Pittsburgh, Pennsylvania, USA. Data collection took place during the summer of 2013. Each appliance type is represented by dozens of different instances of varying make/models. For each appliance, three to six measurements were collected.*" (Gao et al., 2014)

- **Electricity Consumption and Occupancy (ECO):** "*In particular, it contains aggregate electricity consumption data – including real and reactive power for each of the three phases – and plug-level measurements of selected household appliances. The data has been collected at 1 Hz granularity and over a period of 8 months. Furthermore, the data set also contains occupancy information of the monitored households.*" (C. Beckel, W. Kleiminger, R. Cicchetti, T. Staake, S. Santini. 2014)

The table below (Table 1) describes in brief the duration, sampling rate, type and country of each of the above-cited datasets:

Dataset	*Duration*	*Type*	*Country*
REDD (Kolter and Johnson, 2011)	19 days	Residential	US

UK-DALE (Kelly and Knottenbelt, 2015b)	2 years	Residential	UK
PLAID (Gao et al., 2014)	-	Residential	US
ECO (C. Beckel, W. Kleiminger, R. Cicchetti, T. Staake, S. Santini. 2014)	8 months	Residential	Switzerland

Table 1 Datasets in brief

Within the frame of the project Non-Intrusive Load Monitoring for Intelligent Home Energy Management Systems, the team supervised by Professor Dr. A. Ruano worked on a specific case study in a residential home in Algarve, Portugal. One of the project goals is to do energy monitoring using NILM that disaggregates the overall energy usage using the load from the utility service entry. The data was collected from a private residential house, recorded at a sampling rate of 1s.

The structure of the house and the appliances it has are described in the quote bellow: "*It is a detached house, with two floors and with 20 different spaces (including garden, halls, and so on). The house has a PV installation, composed of 20 Sharp NU-AK panels, arranged in two strings, each panel with a maximum power of 300W. The inverter is a Kostal Plenticore Plus converter (KI), which also controls a BYDBattery BoxHVH11.5 (with a storage capacity of 11.5 kWh). Several electrical appliances exist in this house, and a json file was created according to the format used by the NILM Toolkit. The house electric panel is a Schneider panel consisting of 16 monophasic circuit breakers, plus a triphasic one. The house also has available a few TP-Link HS100 Wi-Fi Smart Plugs (SP), one Intelligent Weather Station (IWS), and a few Self-Powered Wireless Sensors (SPWS) for measuring room climate variables*." (A. Ruano, K. Bot, M. Graça Ruano, 2020)

1.2 Deep Learning

Two different Deep Learning models are employed in this work. They are briefly described below.

1.2.1 Convolutional Neural Network

Convolutional Neural Network, also known as CNN or ConvNet, is a deep learning model that specializes in processing data that has a grid-like topology, such as images. Figure 5 describes the general model of a CNN.

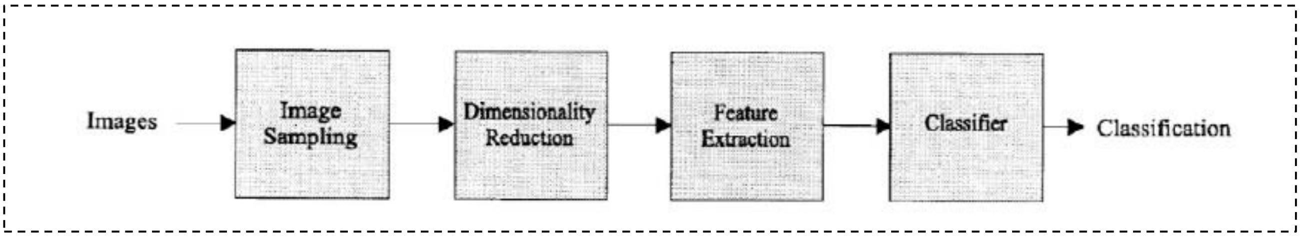

Figure 5 CNN general model (L., S., Giles, C. L., Tsoi, A. C., and Back, A. D. (1997))

A CNN typically has three layers: a convolutional layer, a pooling layer, and a fully connected layer.

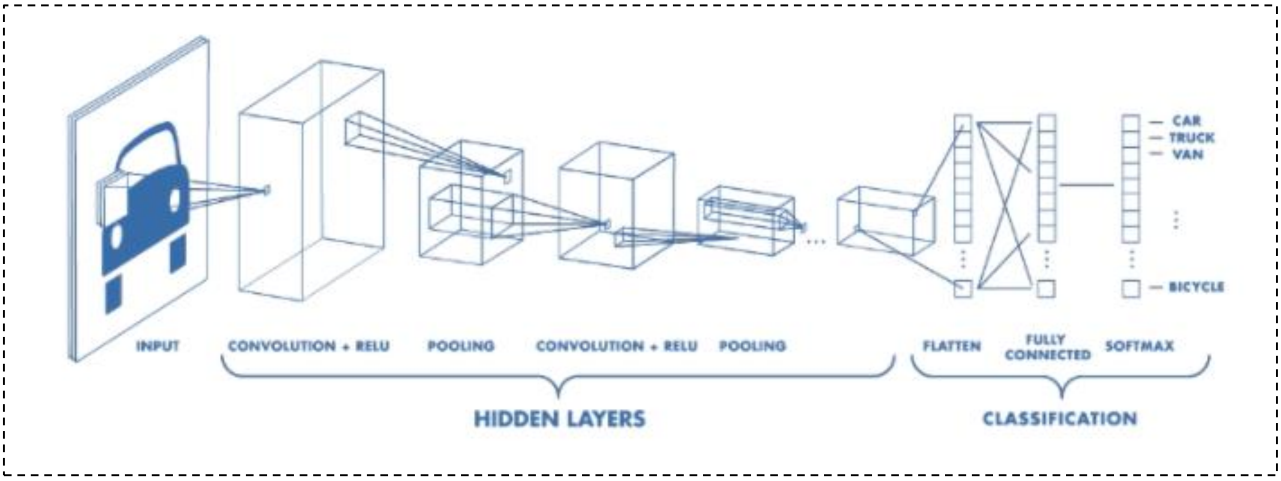

Figure 6 CNN architecture (S. Patel, J. Pingel, 2017)

1 - Convolution Layer: where features are extracted from the input image.

2- Pooling Layer: used to reduce the dimensionality of each feature. It helps to reduce the number of parameters and computations needed. CNN uses max-pooling where it chooses the largest element from the spacial neighborhood defined. (S. Patel, J. Pingel, 2017)

4 - Fully-Connected Layer

The output from the convolution and pooling layers represent high-level features of the input image. The Fully-Connected layers use these features for classifying the input image into various classes based on the training dataset.

1.2.2 Long Short-term Memory Recurrent Neural Network

Long Short Term Memory networks, shortened as "LSTM", are a special kind of recurrent neural networks. As a work of Hochreiter & Schmidhuber in 1997, LSTMs were first introduced to solve the vanishing gradient problem in RNNs. They are capable of learning long-term dependencies by storing processed information about longer sequence of data.

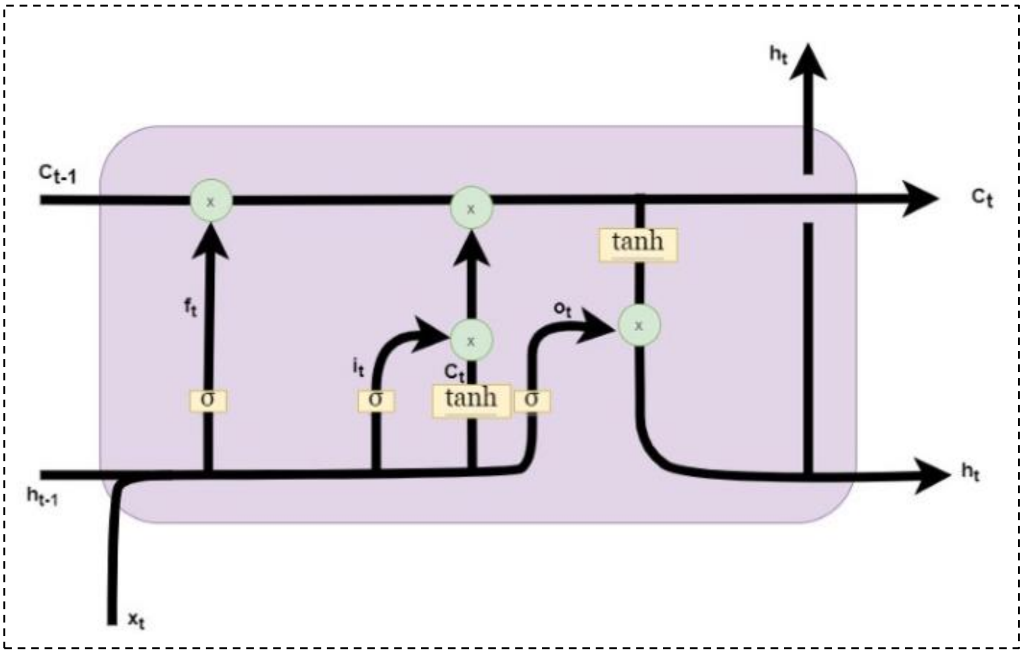

Figure 7 LSTM Structure (M. Phi, 2018)

An LSTM consists of a set of recurrently connected blocks known as memory blocks. Each LSTM is made of one or multiple recurrently connected memory cells as well as the input, the output and forget gates. (Alex Graves, et al.,2005)

LSTM has a similar flow as an RNN. It processes data, sequentially passing on information as it propagates forward; the difference are the operations within the cells.

The forget gate will give an output based on the input feature and the hidden state. The sigmoid is used as an activation function, and the output ranges between 0 and 1. The output will then be multiplied by the cell state, thus resulting in two possible results: if the output is 0 the cell state will be empty, if it is 1 then the cell state will remain the same.

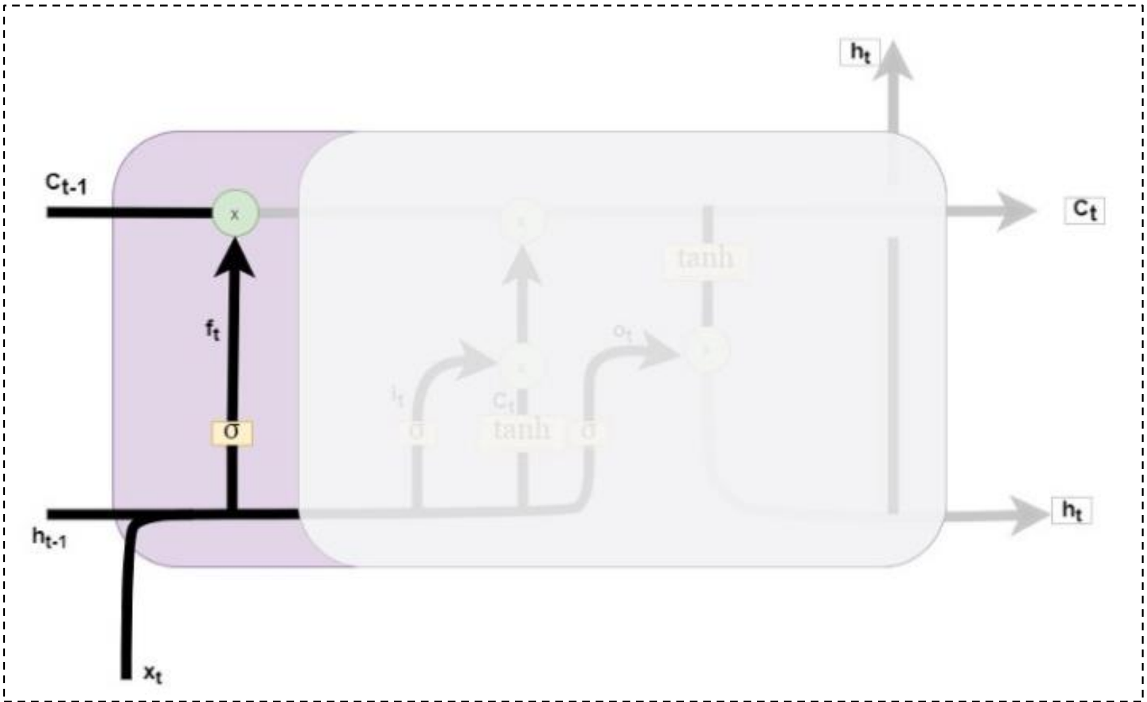

Figure 8 Forget Gate (M. Phi, 2018)

After the forget gate stage, the input gate will be given the hidden state of t-1 and the input feature. The input gate will determine what new data to keep. The “tanh” function creates a new set of values to be stored into the memory. The generated values will be multiplied by the output from the input gate then added to the cell state.

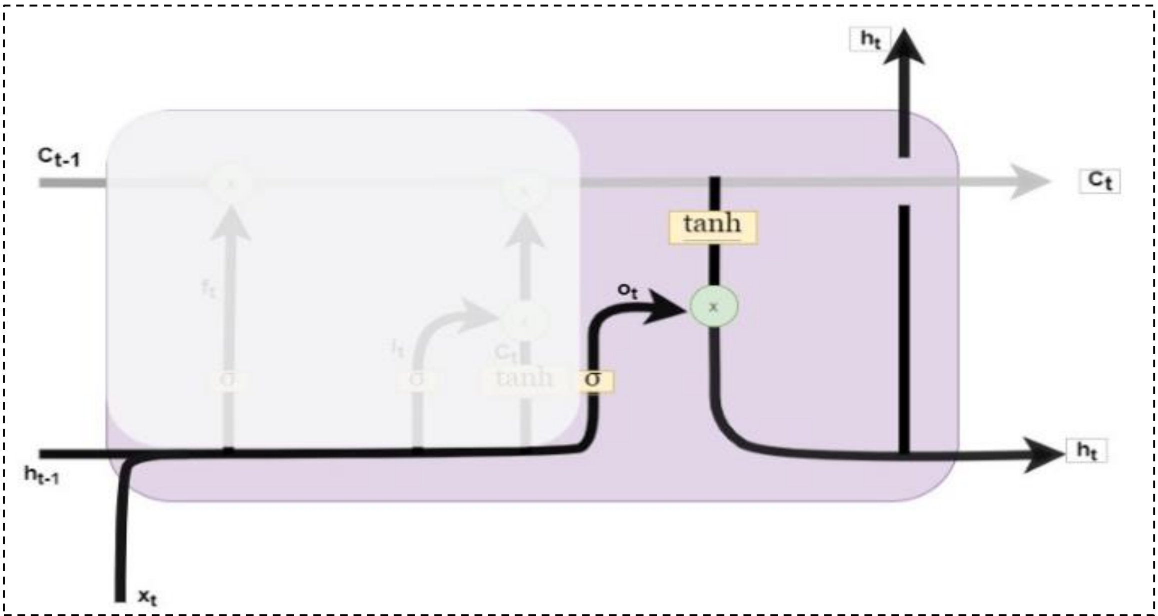

Figure 9 Output Gate (M. Phi, 2018)

After the output gate and the tanh function decide what data should be output and transformed into the range -1 to 1, the resulting new data will be output along with Ct.

1.3 Performance Criteria

The following performance criteria are commonly used in the non-intrusive load monitoring field:

- Mean Absolute Error (MAE): The absolute value of the difference between the predicted and the actual value.
- F1 score: The weighted average of the precision and recall values where the best value reaches 1 and the worst 0.
- Signal Aggregate Error: relative error of the total energy.
- Estimated Accuracy: The degree to which a prediction varies to its actual value.

1.3.1.1 MAE

$$MAE = \frac{\sum_{i=1}^{n} |y_i - x_i|}{\text{n}}$$

Where

y_i: prediction

x_i: true value

n: total number of data points

1.3.1.2 F1 Score

$$F1 = 2 * \frac{precision * recall}{precision + recall}$$

Where

$$Precision = \frac{TP}{TP + FP}$$

And

$$Recall = \frac{TP}{TP + FN}$$

TP: True Positives

FP: False Positives

FN: False Negatives

1.3.1.3 Signal Aggregate Error

$$SAE = \frac{\sum_{i=1}^{n} y_i \; - \; \sum_{i=1}^{n} x_i}{\sum_{i=1}^{n} x_i}$$

Where:

Yi: prediction

Xi: true value

n: total number of data points

1.3.1.4 Estimated Accuracy

$$Estimated\ Accuracy = \frac{\sum_{t=1}^{T} |\ \mathrm{x}_t{}^{(\mathrm{m})} - \mathrm{y}_\mathrm{t}{}^{(\mathrm{m})}|}{2 * \ \sum_{t=1}^{T} \mathrm{y}_\mathrm{t}{}^{(\mathrm{m})}}$$

Where:

T: the time sequence or number of disaggregated readings

$\mathbf{x}_t$: the estimated power consumed at time t for appliance m

$\mathbf{y}_\mathbf{t}$: the ground truth power consumed at time t for appliance m

1.4 Thesis Outline

This thesis is organized according to the following outline. In chapter 1, up to now, we introduced the necessary background for the reader to be comfortable reading the rest of the thesis. In chapter 2, we will discuss some previous work done in non-intrusive load monitoring using deep learning technics. Next in chapter 3, we will detail the proposed work. Then in chapter 4, we will apply the two introduced models, CNN and LSTM, on the two public datasets:

- UKDALE
- REDD

Chapter 5, will discuss the application of the same models on the dataset from the Intelligent Home Energy Management Systems project. To finish, in chapter 6 final conclusions and future work will be drawn.

2 Chapter 2: Previous Work

2.1 State of the art

All the papers were collected from the trusted commonly reliable website, Web of Science. The papers were looked collected using the following keywords:

- NILM
- Non-intrusive Load management using deep learning
- Deep Learning Techniques applied on non-intrusive load management

The first search produced around 50 papers. Then it was narrowed to the 5 papers chosen below based on their suitable titles, interesting abstracts, and their recent publication date of 2020.

Paper 1: Investigation of Deep Learning-based Techniques for Load Disaggregation, Low-Frequency Approach (A. Alkhulaifi, A. J. Aljohani (2020))

In the paper "Investigation of Deep Learning-based Techniques for Load Disaggregation, Low-Frequency Approach", the authors used the following models to disaggregate the energy load:

- Denoising Auto-encoder (DAE)
- Recursive Neural Network (RNN) Long Short-Term Memory networks (LSTM)
- Sequence-to-Point Neural Network (Seq2Point)
- Gated Recurrent Units Recurrent Neural Networks (RNN GRU)

The datasets used in this paper are the REDD and UK-DALE where the chosen appliances for the experiment are the single state (on/off) and multi-state appliances. The models were trained and tested using the two datasets. The metric the authors chose to evaluate the models is the Mean Absolute Error (MAE). The obtained results are shown below (tables 2 to 5); the best results are in bold (A. Alkhulaifi, A. J. Aljohani (2020)):

Appliance	**DAE**	**RNN LSTM**	**Seq2Point**	**RNN GRU**
Microwave	26.39	42.04	**13.15**	34.58
Dish Washer	51.02	90.76	**9.93**	62.77

Table 2 - Appliance MAE, in watts, for REDD data set. (A. Alkhulaifi, A. J. Aljohani (2020))

Appliance	**DAE**	**RNN LSTM**	**Seq2Point**	**RNN GRU**
Microwave	39.61	57.12	**20.21**	46.64
Dish Washer	61.17	93.18	**16.61**	65.73

Table 3 - Appliance MAE, in watts, for UK-DALE data set. (A. Alkhulaifi, A. J. Aljohani (2020))

Appliance	**DAE**	**RNN LSTM**	**Seq2Point**	**RNN GRU**
Microwave	**46.19**	56.96	66.80	59.98
Dish Washer	152.69	**92.94**	100.78	148.35

Table 4 Appliance MAE In Watts. Trained On Redd and Tested On UK-Dale. (A. Alkhulaifi, A. J. Aljohani (2020))

Appliance	**DAE**	**RNN LSTM**	**Seq2Point**	**RNN GRU**
Microwave	49.22	42.14	**41.97**	54.39
Dish Washer	87.35	138.90	**78.60**	88.26

Table 5 Appliance MAE In Watt Trained On UK-Dale and Tested On Redd. (A. Alkhulaifi, A. J. Aljohani (2020))

Paper 2: A convolutional auto encoder-based approach with batch normalization for energy disaggregation (H. Chen, Y. Wang, C. Fan (2020))

In this paper, the authors used the following models to desegregate the energy load:

- Long Short-Term Memory networks (LSTM)
- Convolutional Auto encoder (CAE)
- Convolutional Auto encoder with batch normalization (CAEBN)
- Convolutional Auto encoder with batch normalization and Hill Climbing (CAEBN-HC)

The dataset used in this paper is the REDD dataset. The metrics that the authors chose to evaluate the models are the Mean Absolute Error (MAE) and the Signal Aggregate Error (SAE). The obtained results are shown below in Table 6, and the best results are in bold (H. Chen, Y. Wang, C. Fan (2020)):

Appliance	Error Measurement	LSTM	CAE	CAEBN	CAEBN-HC
Microwave	MAE	21.64	26.95	21.99	**9.594**
	SAE	0.6944	0.3072	0.41	**0.082**
Fridge	MAE	26.17	24.22	11.98	**7.62**
	SAE	0.3999	0.34	0.078	**0.013**

Table 6 Results (H. Chen, Y. Wang, C. Fan (2020))

Paper 3: Non-Intrusive Load Disaggregation by Convolutional Neural Network and Multi-Label Classification (L. Massidda, M. Marrocu, S. Manca (2020))

The authors employed only a CNN to desegregate the energy load.

The dataset used in this paper is the UK-DALE dataset. The metrics that the authors chose to evaluate the models are the mean absolute error and Signal Aggregate Error. The obtained results are shown below in Tables 7 and 8 (L. Massidda, M. Marrocu, S. Manca (2020)):

	Fridge		Dishwasher		Washing Machine	
	Ensemble Mean	90% Interval	Ensemble Mean	90% Interval	Ensemble Mean	90% Interval
Precision	0.875	(0.867, 0.885)	0.942	(0.904, 0.966)	0.975	(0.968, 0.979)
Recall	0.859	(0.846, 0.871)	0.919	(0.890, 0.942)	0.982	(0.977, 0.987)
Accuracy	0.880	(0.878, 0.882)	0.997	(0.995, 0.997)	0.997	(0.996, 0.997)
F1 score	0.867	(0.864, 0.870)	0.930	(0.905, 0.946)	0.978	(0.976, 0.980)
MCC	0.759	(0.755, 0.762)	0.928	(0.903, 0.945)	0.977	(0.974, 0.979)
MAE [W]	15.25	(15.08, 15.47)	20.41	(19.99, 21.00)	41.97	(41.80, 42.21)
SAE	−0.020	(−0.046, 0.002)	−0.042	(−0.082, −0.005)	−0.077	(−0.085, −0.066)

Table 7 Performance the seen case on the UK-DALE dataset. (L. Massidda, M. Marrocu, S. Manca (2020))

	Fridge		Dishwasher		Washing Machine	
	Ensemble Mean	90% Interval	Ensemble Mean	90% Interval	Ensemble Mean	90% Interval
Precision	0.892	(0.883, 0.898)	0.788	(0.738, 0.826)	0.858	(0.811, 0.893)
Recall	0.851	(0.841, 0.861)	0.835	(0.768, 0.897)	0.869	(0.827, 0.918)
Accuracy	0.905	(0.900, 0.908)	0.989	(0.987, 0.990)	0.997	(0.996, 0.998)
F1 score	0.871	(0.863, 0.876)	0.809	(0.790, 0.822)	0.863	(0.835, 0.900)
MCC	0.796	(0.786, 0.803)	0.805	(0.784, 0.817)	0.862	(0.834, 0.899)
MAE [W]	17.03	(16.82, 17.24)	33.07	(31.19, 35.68)	8.31	(7.88, 8.70)
SAE	−0.046	(−0.066, −0.025)	0.063	(−0.054, 0.219)	0.014	(−0.059, 0.115)

Table 8 Performance for the unseen case on the UK-DALE dataset. (L. Massidda, M. Marrocu, S. Manca (2020))

Paper 4: A Practical Solution for Non-Intrusive Type II Load Monitoring Based on Deep Learning and Post-Processing (W. Kong, Z. Y. Dong, B. Wang, J. Zhao, J. Huang (2020))

The authors of this paper also used the Convolutional neural network model to desegregate the energy. The dataset used in this paper is the UK-DALE dataset where the appliances chosen were type 2 appliances. The dataset was preprocessed in the paper using data augmentation. Then it was preprocessed using another CNN that classifies whether a sequence of estimated consumptions belongs to the target appliance or not. The metrics the authors chose to evaluate the models are the Estimated Accuracy and the F1.

The obtained results are shown below in "Table 7" (W. Kong, Z. Y. Dong, B. Wang, J. Zhao, J. Huang (2020)):

F1	House 1	Dishwasher: 0.916 Washing Machine: 0.928
	House 2	Dishwasher: 0.879 Washing Machine: 0.897
	House 3	Dishwasher: 0.854 Washer Dryer: 0.812
Accuracy	House 1	Dishwasher: 0.895 Washing Machine: 0.920
	House 2	Dishwasher: 0.959 Washing Machine: 0.842
	House 3	Dishwasher: 0.881 Washer Dryer: 0.735

Table 9 Paper 4 results (W. Kong, Z. Y. Dong, B. Wang, J. Zhao, J. Huang (2020))

Paper 5: A Deep Recurrent Neural Network for Non-Intrusive Load Monitoring Based on Multi-Feature Input Space and Post-Processing (H. Rafiq, X. Shi, H. Zhang, H. Li, M. K. Ochani (2020))

In the paper "A Deep Recurrent Neural Network for Non-Intrusive Load Monitoring Based on Multi-Feature Input Space and Post-Processing", the authors used the Deep LSTM model to desegregate the energy. The datasets used in this paper are the UK-DALE dataset and the Electricity Consumption and Occupancy (ECO) dataset where the appliance chosen were type 1 and 2 appliances. The dataset was post-processed in the paper where irrelevant activations were eliminated during the disaggregation stage by comparing the lengths of ground-truth and predicted appliance activations of both type-1 and type-2 appliances. The metrics the authors chose to evaluate the models are the Precision, Recall, and F1, mean absolute error (MAE), signal aggregate error (SAE), and estimation accuracy (EA).

The obtained results are shown below in "Tables 12 to 14" (H. Rafiq, X. Shi, H. Zhang, H. Li, M. K. Ochani (2020)):

House #	Appliances	Without Post-Processing				With Post-Processing			
		F1	MAE (W)	SAE	EA	F1	MAE (W)	SAE	EA
1	Kettle	0.658	2.162	0.179	0.911	0.995	1.837	0.217	0.891
1	Fridge	0.497	13.980	0.138	0.919	0.997	5.679	0.347	0.826
5	Microwave	0.535	65.090	0.028	0.986	0.719	21.450	0.515	0.743
2	Dishwasher	0.559	13.076	0.094	0.953	0.749	5.877	0.419	0.790
1	Washing Machine	0.322	65.720	1.010	0.492	0.795	18.870	0.655	0.673
2	Electric Stove	0.886	3.519	0.623	0.688	0.981	0.240	0.005	0.997
2	Television	0.976	0.976	0.018	0.991	0.995	0.497	0.012	0.994
	Overall	**0.633**	**23.503**	**0.298**	**0.848**	**0.890**	**7.778**	**0.310**	**0.845**

Table 10 Performance on a seen UK-DALE. (H. Rafiq, X. Shi, H. Zhang, H. Li, M. K. Ochani (2020))

House #	Appliances	Without Post-Processing				With Post-Processing			
		F1	MAE (W)	SAE	EA	F1	MAE (W)	SAE	EA
5	Kettle	0.701	14.973	0.685	0.657	0.965	1.966	0.058	0.971
5	Fridge	0.732	27.863	0.270	0.865	0.872	19.608	0.467	0.766
5	Microwave	0.242	0.546	0.504	0.748	0.317	0.392	0.828	0.586
5	Dishwasher	0.554	35.129	0.273	0.863	0.809	15.275	0.323	0.838
5	Washing Machine	0.189	30.990	2.18	−0.09	0.765	14.422	0.512	0.744
	Overall	**0.484**	**21.900**	**0.782**	**0.609**	**0.746**	**10.333**	**0.438**	**0.781**

Table 11 Performance on unseen UKDALE. (H. Rafiq, X. Shi, H. Zhang, H. Li, M. K. Ochani (2020))

House #	Appliances	Without Post-Processing				With Post-Processing			
		F1	MAE (W)	SAE	EA	F1	MAE (W)	SAE	EA
2	Kettle	0.961	3.906	0.004	0.998	0.981	2.353	0.043	0.978
2	Fridge	0.838	13.667	0.170	0.915	0.995	4.039	0.121	0.939
2	Microwave	0.721	7.285	0.276	0.862	0.869	5.402	0.437	0.781
2	Dishwasher	0.745	25.736	0.024	0.988	0.891	12.346	0.288	0.856
2	Washing Machine	0.189	30.990	0.686	−0.09	0.701	5.400	0.641	0.679
2	Rice Cooker	0.299	8.900	0.699	−0.161	0.781	1.115	0.378	0.811
5	Electric Oven	0.550	68.611	0.448	0.594	0.736	28.911	0.013	0.993
5	Television	0.512	5.695	0.219	0.890	0.879	3.428	0.649	0.675
	Overall	**0.688**	**23.541**	**0.361**	**0.714**	**0.976**	**8.999**	**0.367**	**0.959**

Table 12 Performance on the ECO datasets. (H. Rafiq, X. Shi, H. Zhang, H. Li, M. K. Ochani (2020))

2.1.1 Results Comparison

This section presents a comparison between the considered papers. The best values are in bold. The comparison is made based on the results found in the papers of two metrics cited above:

- Mean Absolute Error (MAE)
- F1 Score

Acronyms used in the tables below:

MW: Microwave
WM: Washing Machine
DW: Dishwasher
WD: Washer Drier

The table below compares the papers based on **Mean Absolute Error (MAE)**:

Model / Paper/ dataset		**DAE**	**LSTM**	**Seq2Point**	**RNN GRU**	**CAE**	**CAE with batch normalization**	**CNN**
1	**REDD**	MW:26.39 DW:51.02	MW:42.04 DW:90.76	MW:**13.15** DW:**9.93**	MW:34.58 DW:62.77	-	-	-
	UK-DALE	MW:39.61 DW:61.17	MW:57.12 DW:93.18	MW:20.21 DW:16.61	MW:46.64 DW:65.73	-	-	-
2	**REDD**	-	-	-	-	MW:26.95 Fridge:24.22	MW:21.99 Fridge:11.98	
	UK-DALE	-	-	-	-	-	-	-
3	**REDD**	-	-	-	-	-	-	-

	UK-DALE	-	-	-	-	-	-	*(seen):* WM: 41.97 DW: 20.41 Fridge: 15.25 *(unseen):* WM: 8.31 DW: 33.07 Fridge: 17.03
5	**UK-DALE**	-	*(seen):* MW: 21.45 DW: **5.877** Fridge: **5.679** *(unseen):* MW: **0.392** DW: 15.275 Fridge: 19.608	-	-	-	-	-
	ECO	-	MW: **5.402** DW: **12.346** Fridge: **4.039**	-	-	-	-	-

Table 13 MAE Results

The table below compares the papers based on **F1 Score**:

Model / **Paper/dataset**		**CNN**	**LSTM**
3	**UK-DALE**	*(seen):* WM: **0.98** DW: **0.93** Fridge: 0.87 *(unseen):* WM: 0.86 DW: 0.81 Fridge: 0.87	-
4	**UK-DALE**	House 1: DW: 0.916 WM: 0.928	-

<table>
<tr><td></td><td></td><td><table><tr><td>House 2</td><td>DW: 0.879
WM: 0.897</td></tr><tr><td>House 3</td><td>DW: 0.854
WD: 0.812</td></tr></table></td><td></td></tr>
<tr><td rowspan="2">5</td><td>UK-DALE</td><td>-</td><td>(seen):
MW: 0.719
DW: 0.749
Fridge: 0.997
(unseen):
MW: 0.317
DW: 0.809
Fridge: 0.872</td></tr>
<tr><td>ECO</td><td>-</td><td>MW: 0.869
DW: 0.891
Fridge: 0.995</td></tr>
</table>

From the cited comparison, we can note that the best values obtained are as follows:

- **UK-DALE:**

	MAE	**F1**
Fridge	5.679 (Using LSTM)	0.997 (Using LSTM)
Dishwasher	5.877 (Using LSTM)	0.93 (Using CNN)
Microwave	0.392 (Using LSTM)	0.98 (Using CNN)

- **ECO:**

	MAE	**F1**
Fridge	4.039 (Using LSTM)	0.995 (Using LSTM)
Dishwasher	12.346 (Using LSTM)	0.891 (Using LSTM)
Microwave	5.402 (Using LSTM)	0.869 (Using LSTM)

- **REDD:**

	MAE	**F1**
Fridge	-	
Dishwasher	9.93 (Using Seq2Point)	-
Microwave	13.15 (Using Seq2Point)	-

3 Chapter 3: Proposed Work

3.1 Datasets

This thesis will be working with three different datasets:

- UK Domestic Appliance-Level Electricity (UKDale) dataset
- Reference Energy Disaggregation Data set (REDD)
- Intelligent Home Energy Management Systems (IHEMS) Project Data

3.2 Metrics

The metrics used in evaluating the models are two of the widely used metrics in the literature:

- F1 score
- Estimated Accuracy

The detailed formulas can be found above in the section “Performance Criteria”

3.3 Models

Two models will be trained using the datasets described above:

- Convolutional Neural Network
- Long Short-term Memory Recurrent Neural Network

Below are all the parameters used in the models.

The CNN used was introduced in the paper “A Practical Solution for Non-Intrusive Type II Load Monitoring Based on Deep Learning and Post-Processing” by Weicong Kong, Zhao Yang Dong, Bo Wang, Junhua Zhao, Jie Huang. (W. Kong, Z. Y. Dong, B. Wang, J. Zhao, J. Huang (2020))

The LSTM model was built from scratch layer by layer.

3.3.1 CNN Structure

Layer Index	*Layer*
1	Input
2	Conv1D(16,1,relu)
3	Conv1D(16,1,relu)
4	Maxpooling1D
5	Conv1D(32,1,relu)
6	Conv1D(32,1,relu)
7	Maxpooling1D
8	Conv1D(64,1,relu)
9	Conv1D(64,1,relu)
10	Conv1D(64,1,relu)
11	Maxpooling1D
12	Conv1D(128,1,relu)
13	Conv1D(128,1,relu)
14	Conv1D(128,1,relu)
15	Maxpooling1D
16	Dense(1024, relu)
17	Dense(1024, relu)
18	Dense(1, linear)

3.3.2 LSTM Structure

Layer Index	*Layer*
1	Input
2	LSTM(100, tanh)
3	Dropout(0.5)
4	LSTM(100, tanh)
5	Dropout(0.5)
6	Dense(100, relu)
7	Dense(1, linear)

3.4 Technologies Used

Technologies
Jupyter Notebook
Python
Anaconda

3.5 Libraries Used

Thanks to many researchers and developers in the domain, our work was built on their developed build-in libraries. This allowed the work to be smooth and efficient.

Below are the main libraries used in this thesis:

3.5.1 Numpy

"Provides a high-performance multidimensional array and basic tools to compute with and manipulate these arrays." (Conda documentation,2020)

3.5.2 Pandas

"Provides "high-performance, easy-to-use data structures and data analysis tools." pandas provide several methods for reading data in different formats" (Conda documentation, 2020)

3.5.3 Keras

"Allows users to productize deep models on smartphones (iOS and Android), on the web, or on the Java Virtual Machine. It also allows use of distributed training of deep-learning models on clusters of Graphics processing units (GPU) and tensor processing units (TPU)." (Conda documentation,2020)

3.5.4 NilmTK

"Non-Intrusive Load Monitoring Toolkit (NILMTK); an open source toolkit designed specifically to enable the comparison of energy disaggregation algorithms in a reproducible manner. This work is the first research to compare multiple disaggregation approaches across multiple publicly available data sets.

NILMTK includes:

- parsers for a range of existing data sets (8 and counting)
- a collection of preprocessing algorithms
- a set of statistics for describing data sets
- a number of reference benchmark disaggregation algorithms
- a common set of accuracy metrics" (Nilmtk documentation, 2021)

3.6 Experimental Setup

3.6.1 Data Preparation

- Fridge, type 2 appliance, and is the main target of this thesis
- Only house 1 was used for both training and testing of the models in the 2 public datasets UKDale and REDD
- For UKDale, all power load series were down sampled from 6-seconds to 1-minute frequency
- For REDD, all power series were sampled at a 10s sampling rate
- For IHEM Data, all power series were sampled at a 1s sampling rate

In this research the training sets consists of 80% of the dataset and the test set is the remaining 20%. The table below describes the dates windows chosen for the public datasets: UKDale and REDD:

3.6.1.1 UKDALE

Dataset	*Start*	*End*
Training dataset	2014-04-25	2016-04-14
Testing dataset	2016-04-14	2016-06-11

3.6.1.2 REDD

Dataset	*Start*	*End*
Training dataset	2011-04-18	2011-05-20
Testing dataset	2011-05-21	2011-05-24

3.6.2 Data Preprocessing

The dataset was preprocessed using the sliding window technique.

The concept behind this technique as shown in figure 15 is extracting multiple overlapping samples from the sequence inputted.

The Sliding Window technique generates new samples from the existing data, in order to increase the size of the dataset. This method allows the augmentation of the dataset which means the generation of a larger dataset from the existing one. Data augmentation can help avoiding overfitting when training the model.

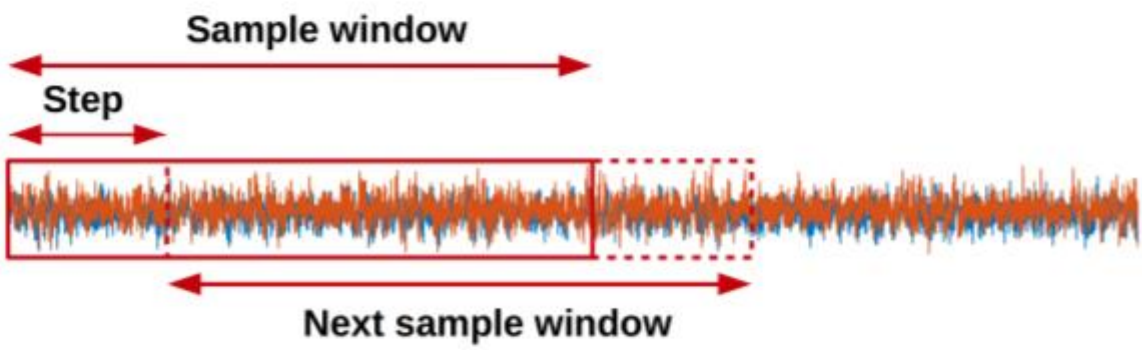

Figure 10 Sliding Window (A. Shenfield, M. Howarth, 2020)

3.7 Summary

The different datasets were set into different time windows then divided into a proportion of 80% to training and 20% to testing to different. Then a data augmentation technique called sliding window was applied to the data. After that, the data was fed to the models described earlier to train the models. Results will be described in details in the upcoming chapters.

4 Chapter 4: Applications using Public Datasets

4.1 UK-Dale

4.1.1 Experimental Setup

The data used from the UKDale dataset was from House 1 and the target appliance was the Fridge. The sampling rate for both the aggregated power and the target power was 60 seconds. Below are some graphs visualizing both the aggregated power and the target appliance power.

Below is the aggregated power plot where the x-axis represents the power consumption in Watt and the y-axis represents time:

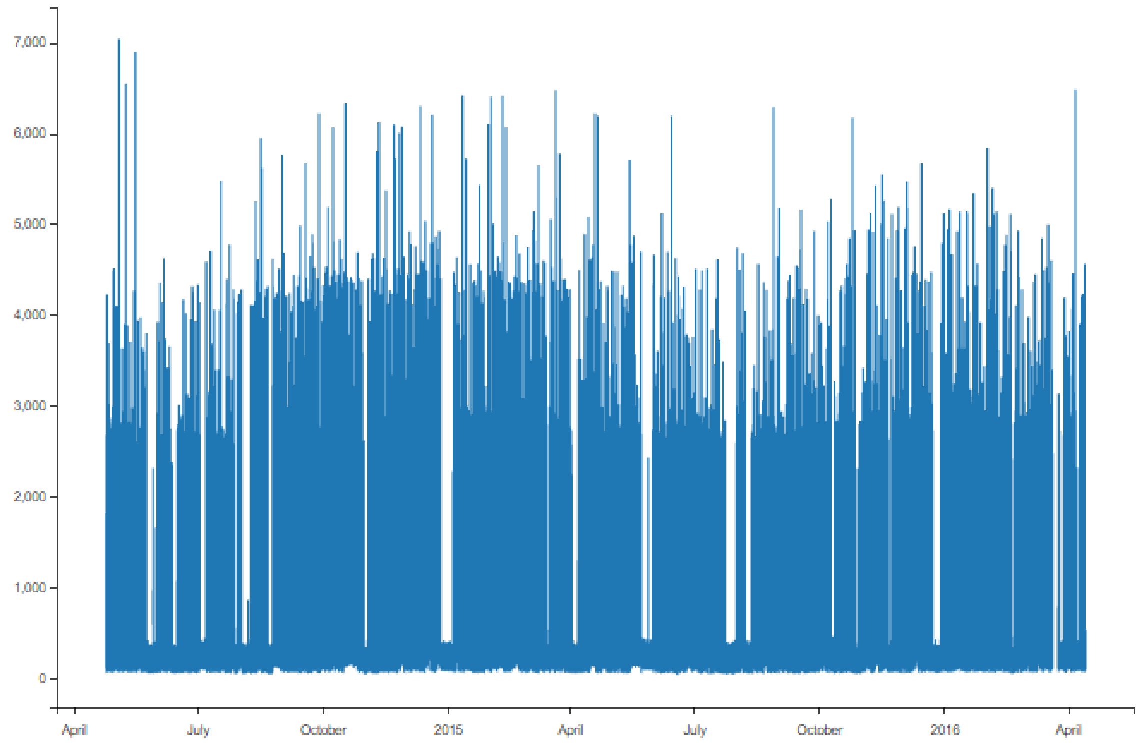

Figure 11 UKDale Aggregated Power

A closer look into the aggregated power graph:

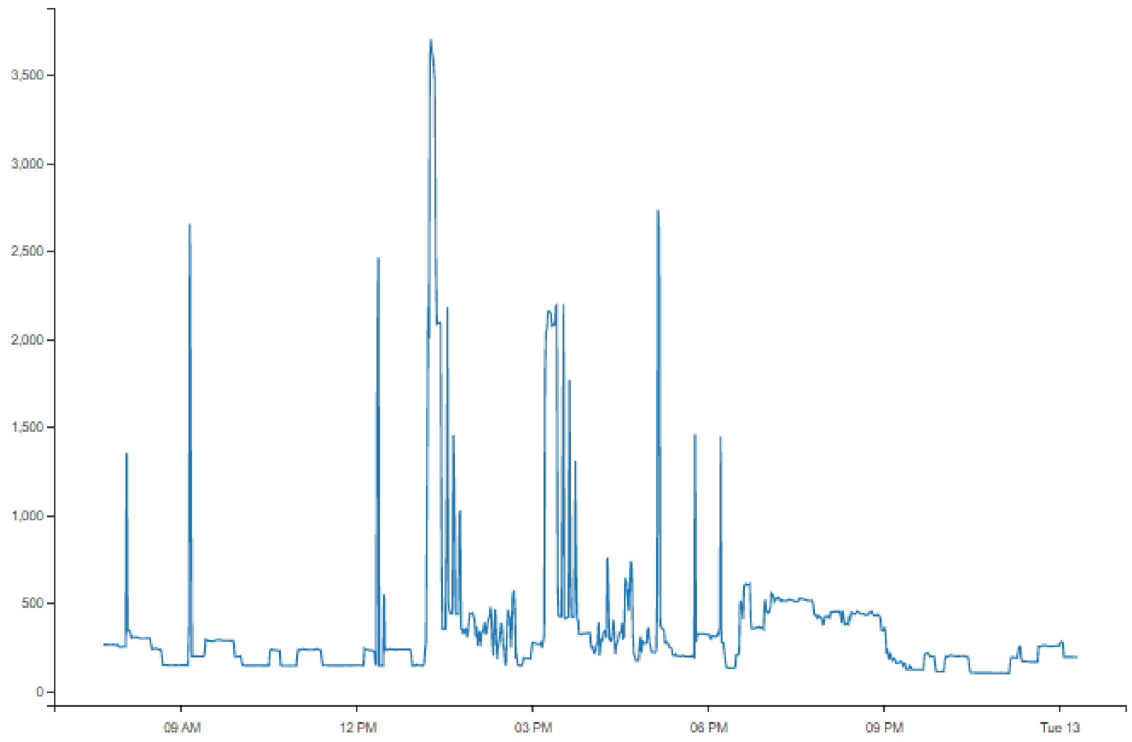

Figure 12 Zoom into the UKDale aggregated power

The target Appliance Power in our case is the Fridge:

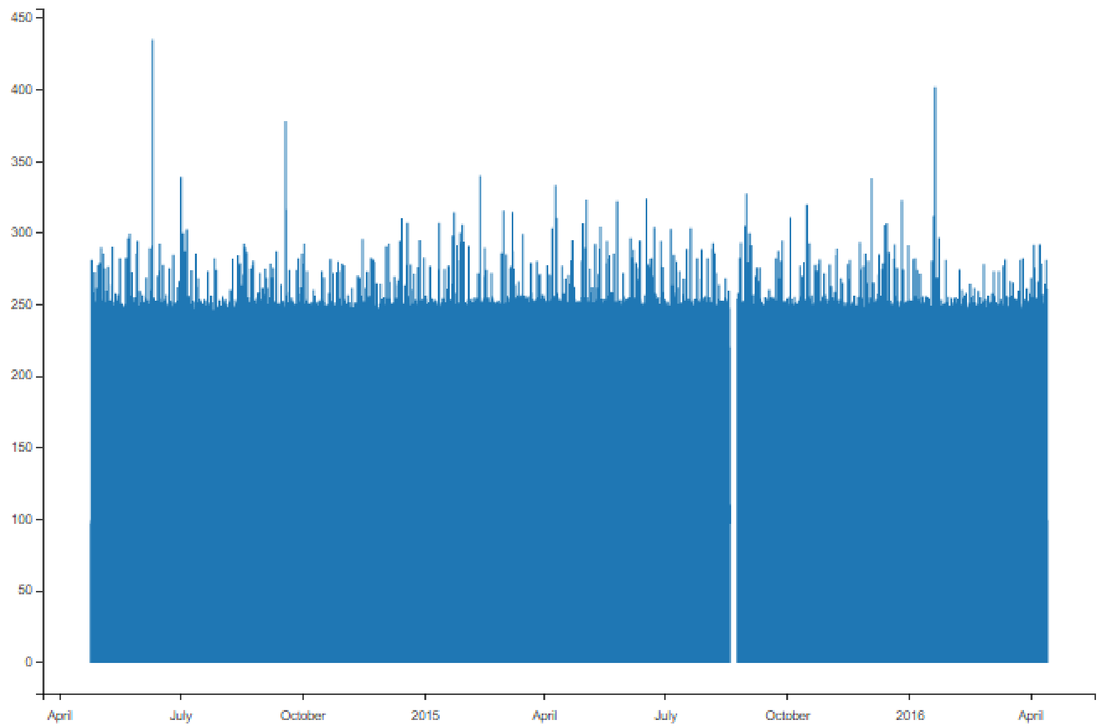

Figure 13 Target Appliance Power

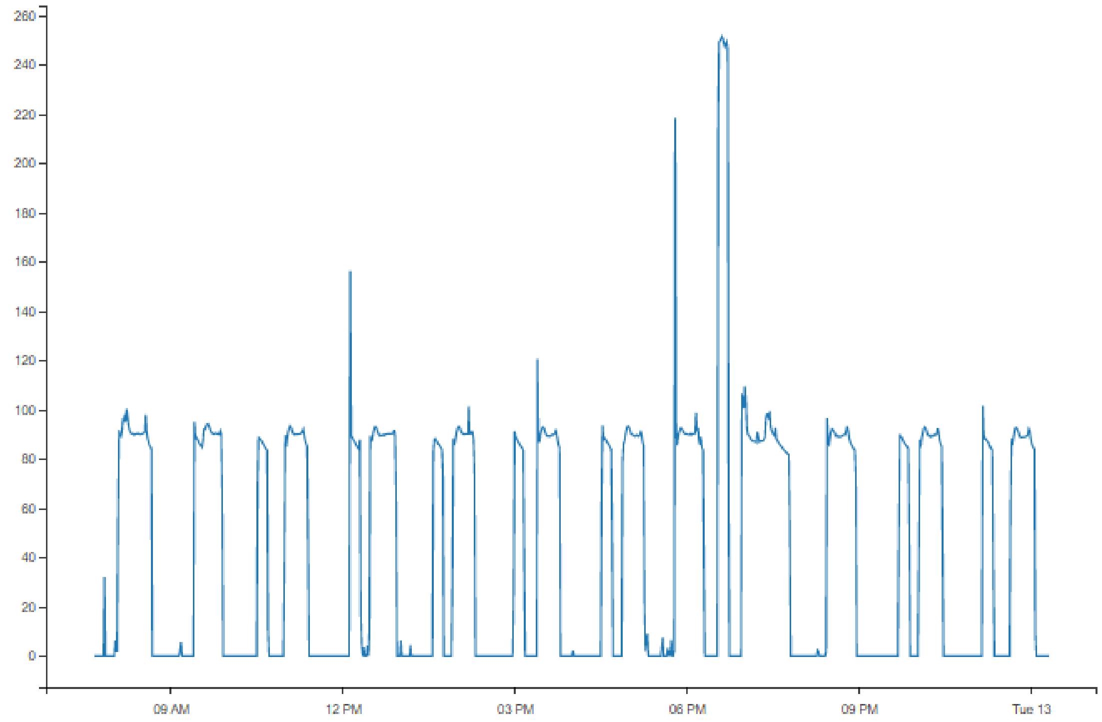

Figure 14 Zoom into the Target appliance power

The chosen dates windows were:

Dataset	*Start*	*End*
Training dataset	2014-04-25	2016-04-14
Testing dataset	2016-04-14	2016-06-11

Some parameters used in our experiment are described in the following table

Parameter	*Value*
Sliding Window	30
Epoch	1000
Batch size	128
Early Stopping	Patience= 200

4.1.2 Results

4.1.2.1 CNN

After the training of the proposed CNN model, the predicted power of the target appliance obtained can be visualized as follows:

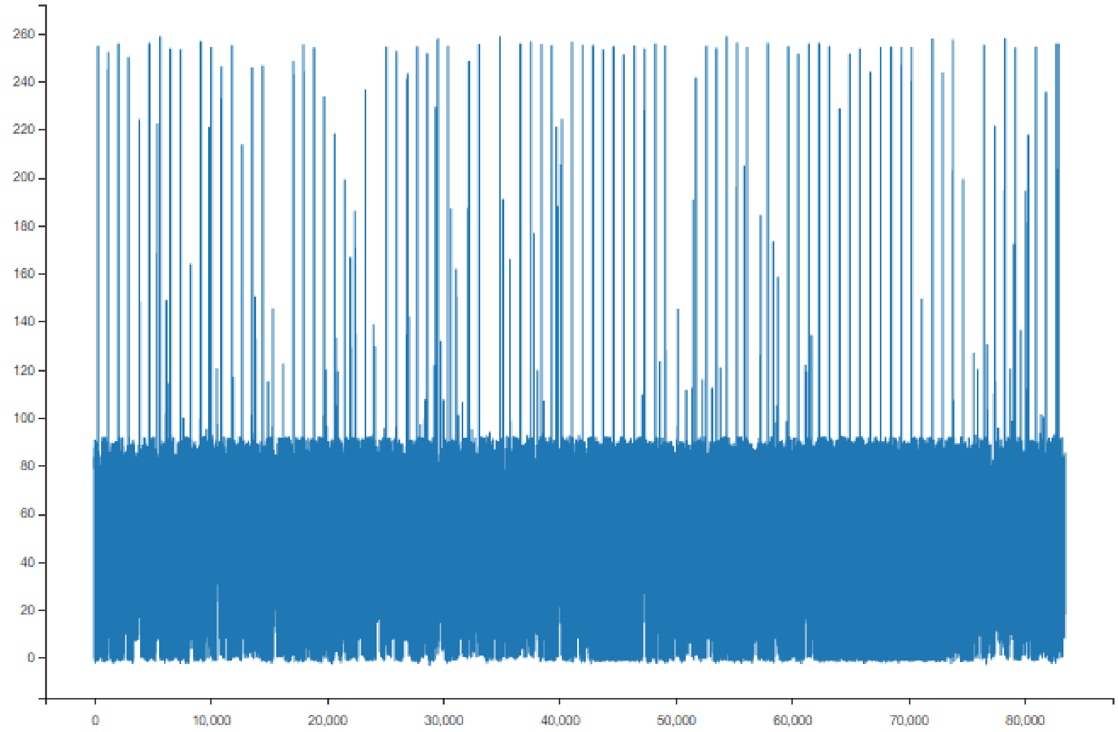

Figure 15 Predicted Power of target appliance

Plotting the predicted power against the testing target appliance power we get the following plot:

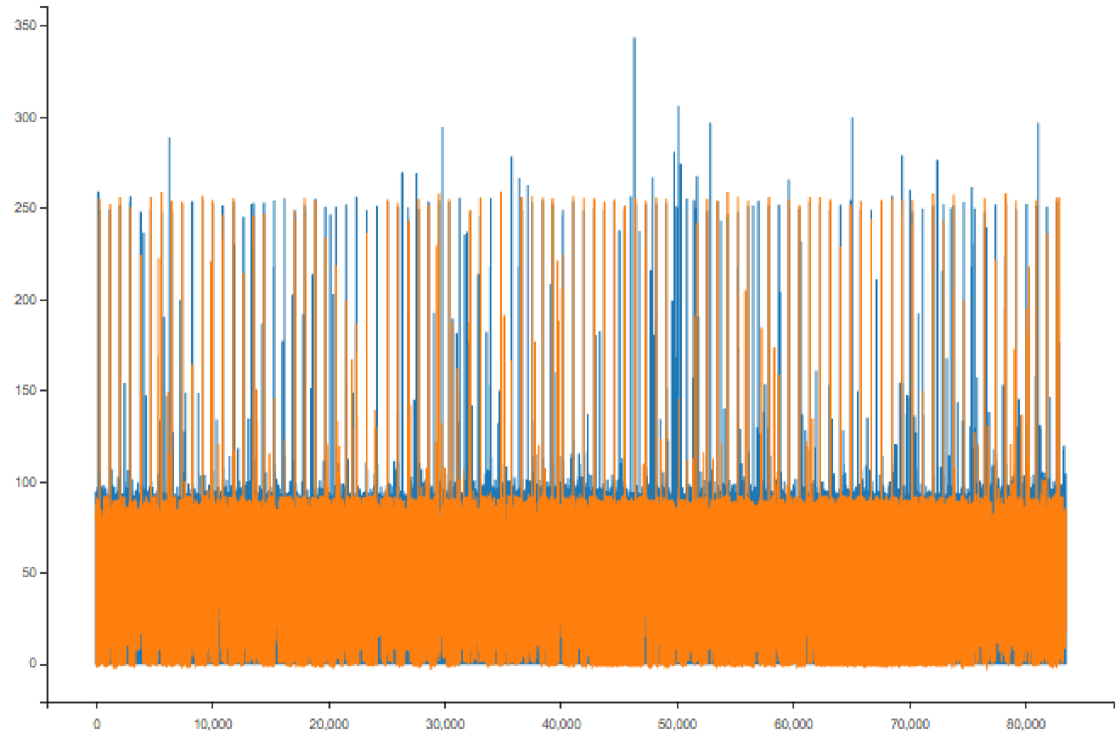

Figure 16 Predicted target power Vs test target power

If we zoom a little deeper we get:

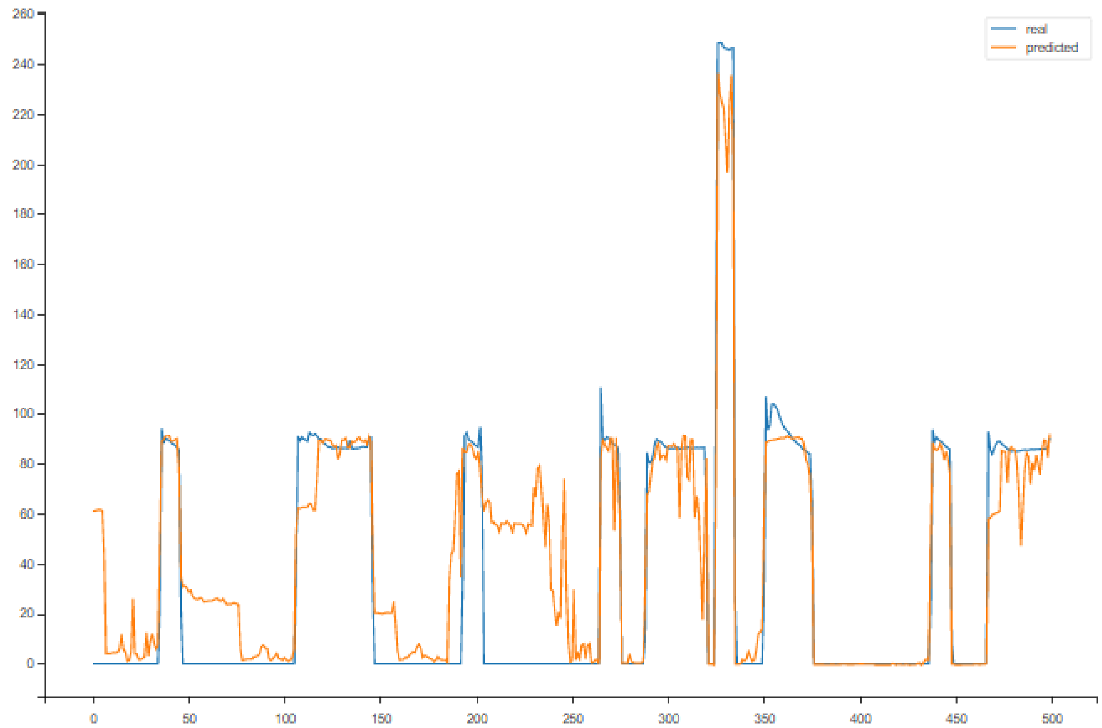

Figure 17 Zoomed Predicted target power Vs test target power

In terms of metrics, the results obtained are as the following table describes:

True Positives	26921
False Positives	4639
True Negatives	45667
False Negatives	6262
Recall	81%
Precision	85%
F1 Score	83%
Estimated Accuracy	78%

4.1.2.2 LSTM

After the training of the proposed LSTM model, the predicted power of the target appliance obtained can be visualized as follows:

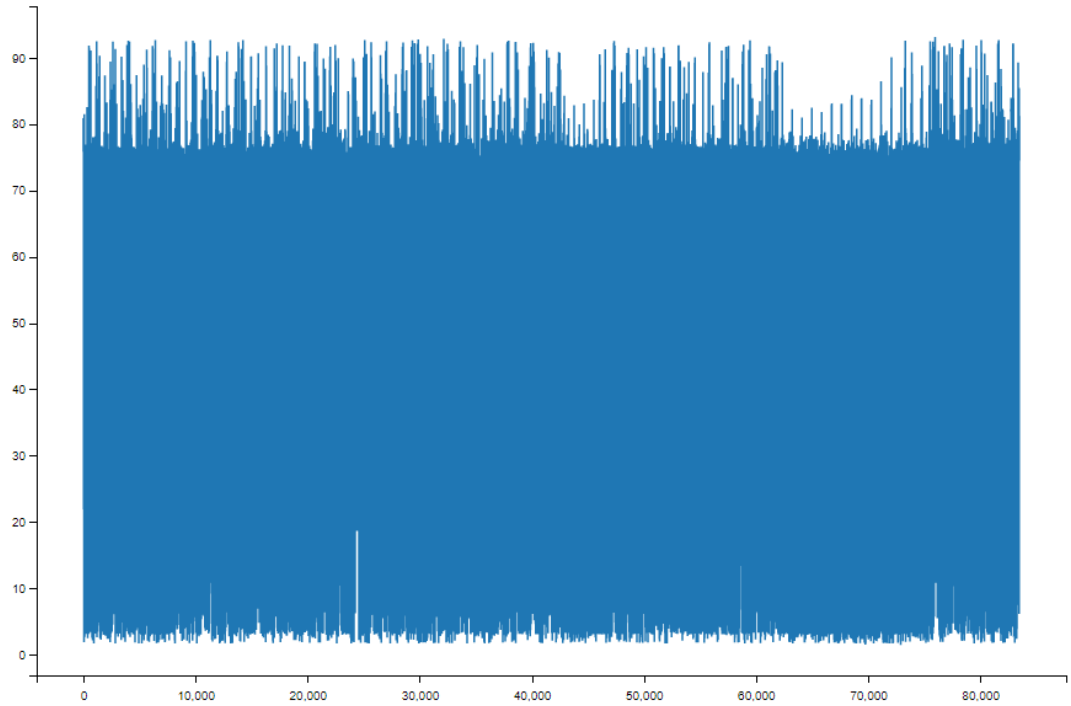

Figure 18 Predicted Target Power LSTM

Plotting the predicted power against the testing target appliance power we get the following plot:

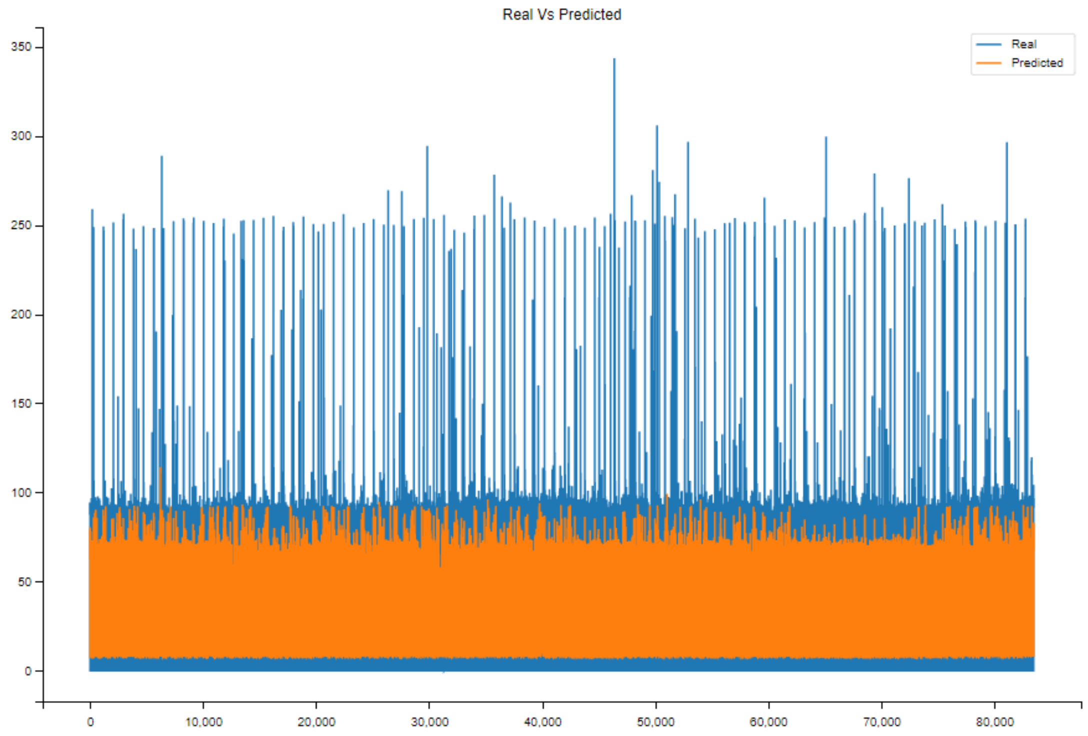

Figure 19 Predicted target power Vs test target power

If we zoom a little deeper we get:

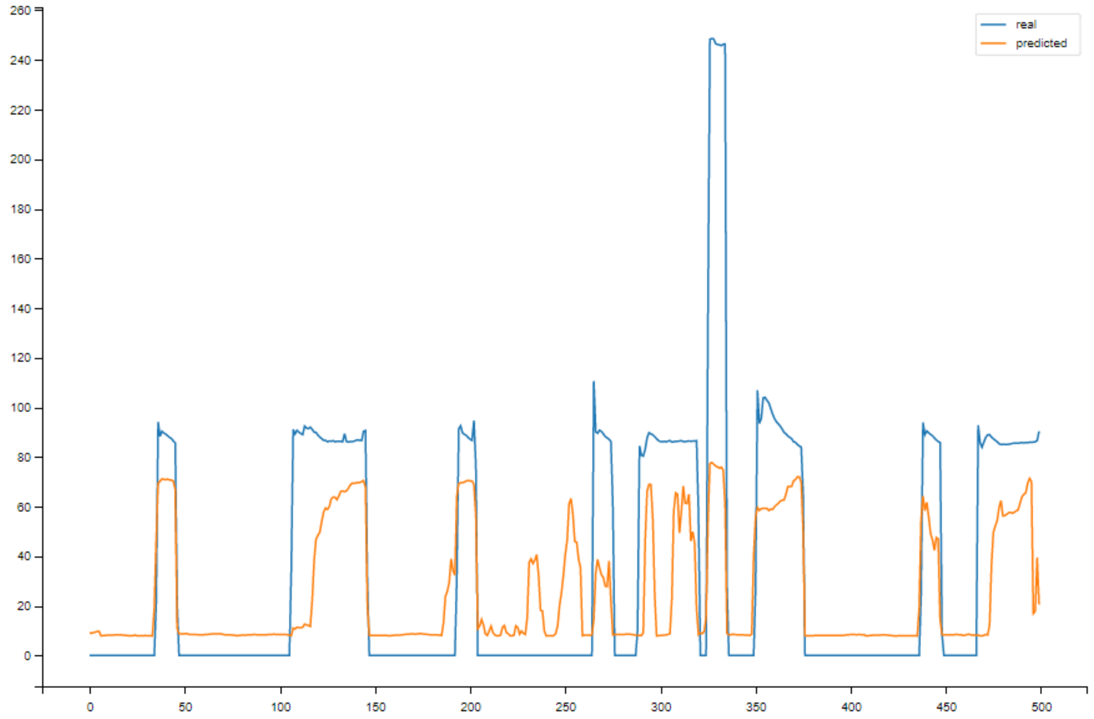

Figure 20 Zoomed Predicted target power Vs test target power

In terms of metrics, the results obtained are as the following table describes:

True Positives	24893
False Positives	5348
True Negatives	44958
False Negatives	8290
Recall	75%
Precision	83%
F1 Score	79%
Estimated Accuracy	69%

4.1.3 Conclusion

The results of the proposed CNN model outperformed the proposed LSTM model.

The results, carried out in the UKDale dataset, in term of F1 score were 83% for the CNN and 79% for the LSTM. Furthermore, in terms of estimated accuracy, the results were 78% for CNN and 69% for the LSTM. In terms of True Positives, the CNN model had 4639 and the LSTM model had 24893.

The following table compares both proposed models CNN and LSTM:

	CNN	*LSTM*
True Positives	26921	24893
False Positives	4639	5348
True Negatives	45667	44958
False Negatives	6262	8290
Recall	81%	75%
Precision	85%	83%
F1 Score	83%	79%
Estimated Accuracy	78%	69%

The proposed method using the CNN model performance compared with the one proposed in the paper (L. Massidda, M. Marrocu, S. Manca (2020)) with no post processing in terms of F1 Score:

	Proposed CNN	*CNN from Massida et. al.*
F1 Score	83%	86%

The proposed method using the LSTM model performance compared with the one proposed in the paper (H. Rafiq, X. Shi, H. Zhang, H. Li, M. K. Ochani (2020)) in terms of F1 Score:

	Proposed LSTM	*LSTM from Rafiq et. al.*
F1 Score	79%	50%

4.2 REDD

4.2.1 Experimental Setup

The data used from the REDD dataset was from House 1 and the target appliance was the Fridge. The sampling rate for both the aggregated power and the target power was 10 seconds. Below are some graphs visualizing both the aggregated power and the target appliance power.

Below is the aggregated power plot where the x-axis represents the power consumption in Watt and the y-axis represents time:

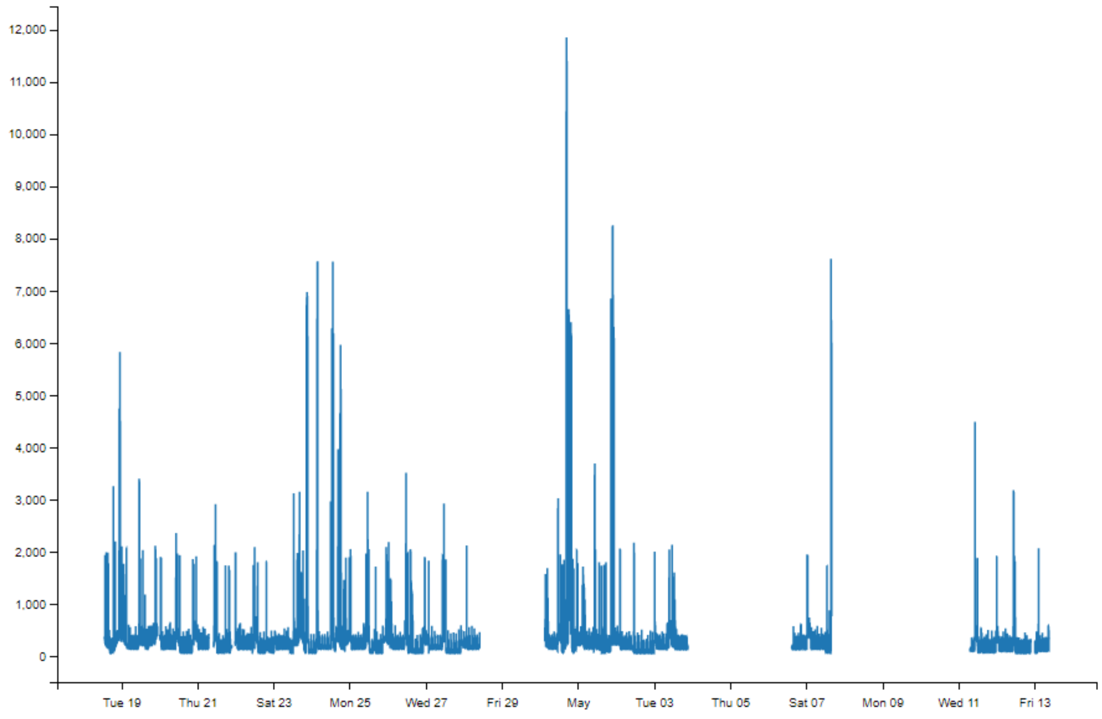

Figure 21 REDD Aggregated Power

A closer look into the aggregated power graph:

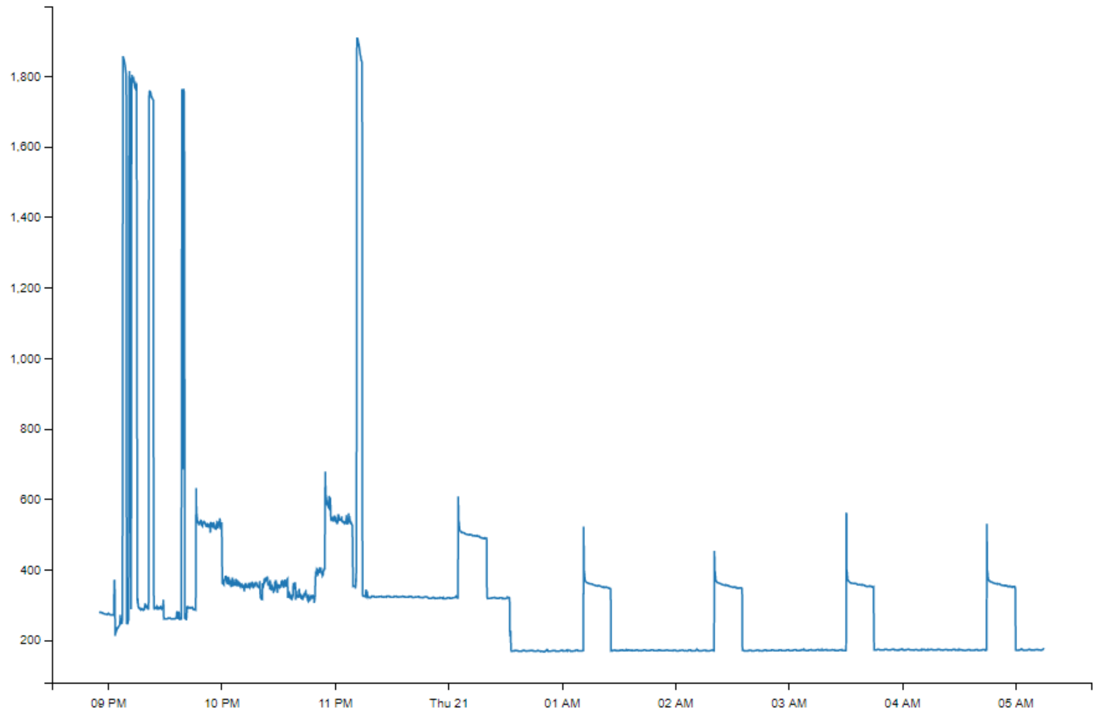

Figure 22 Zoom into the REDD aggregated power

Target Appliance Power in our case is the Fridge:

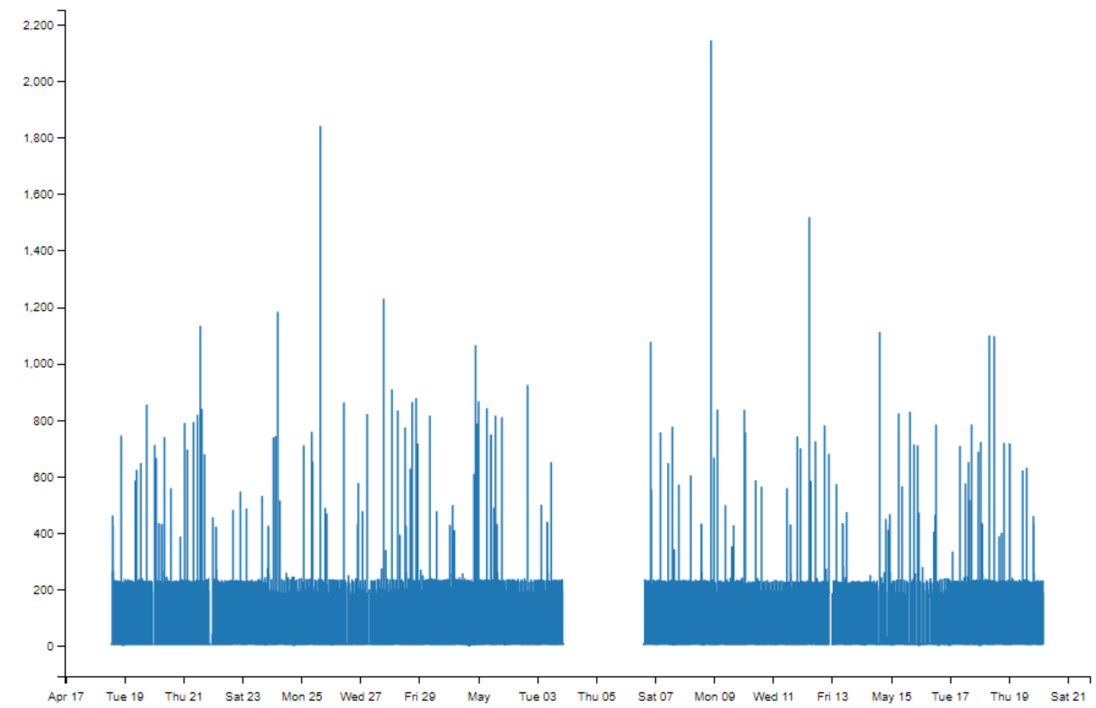

Figure 23 Target Appliance Power

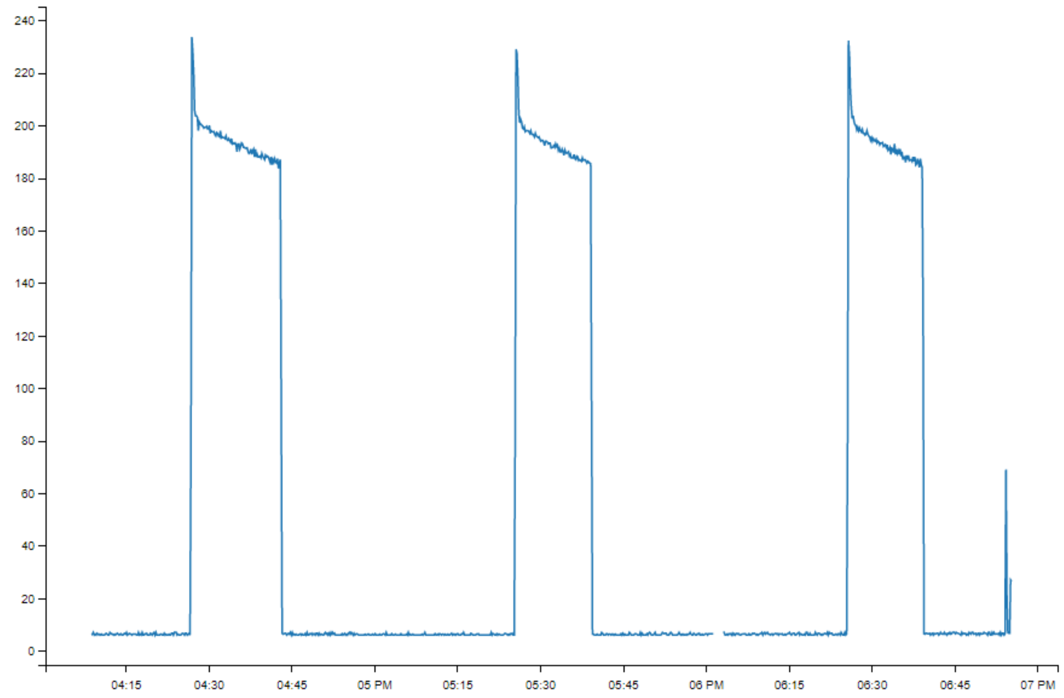

Figure 24 Zoom into the Target appliance power

As we can notice from the graphs above the data has several gaps and NULL values. Because of that we will clean the data by dropping the Nulls. Below are both the aggregated and target power graphs after Null values dropping.

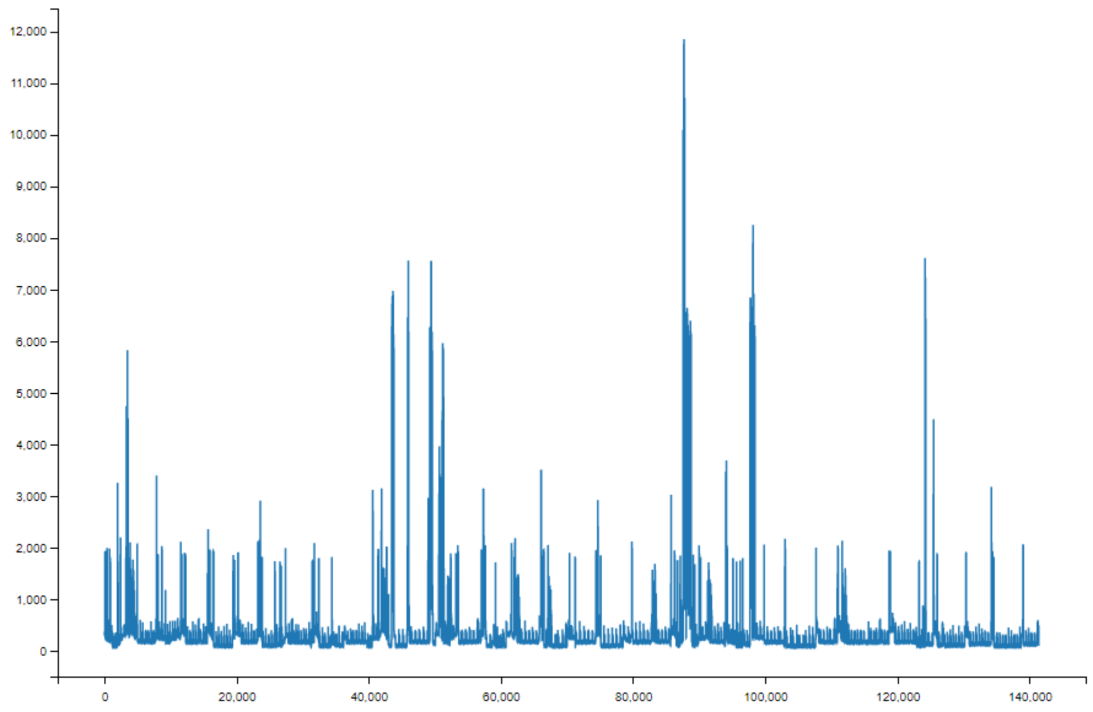

Figure 25 Aggregated power after Null dropping

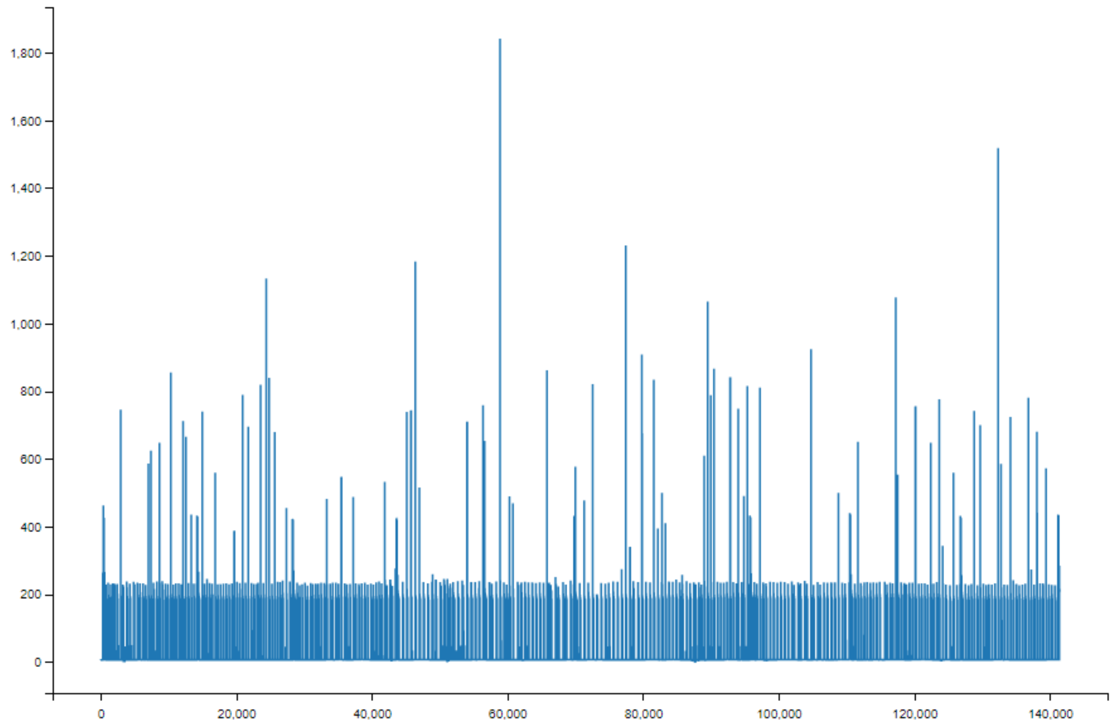

Figure 26 Target power after dropping Nulls

The chosen dates windows were:

Dataset	*Start*	*End*
Training dataset	2011-04-18	2011-05-20
Testing dataset	2011-05-21	2011-05-24

The parameters used in this experiment are:

Parameter	*Value*
Sliding Window	30
Epoch	1000
Batch size	128
Early Stopping	Patience= 200

5.1.1 Results

4.2.1.1 CNN

After the training of the proposed CNN model, the predicted power of the target appliance obtained can be visualized as follows:

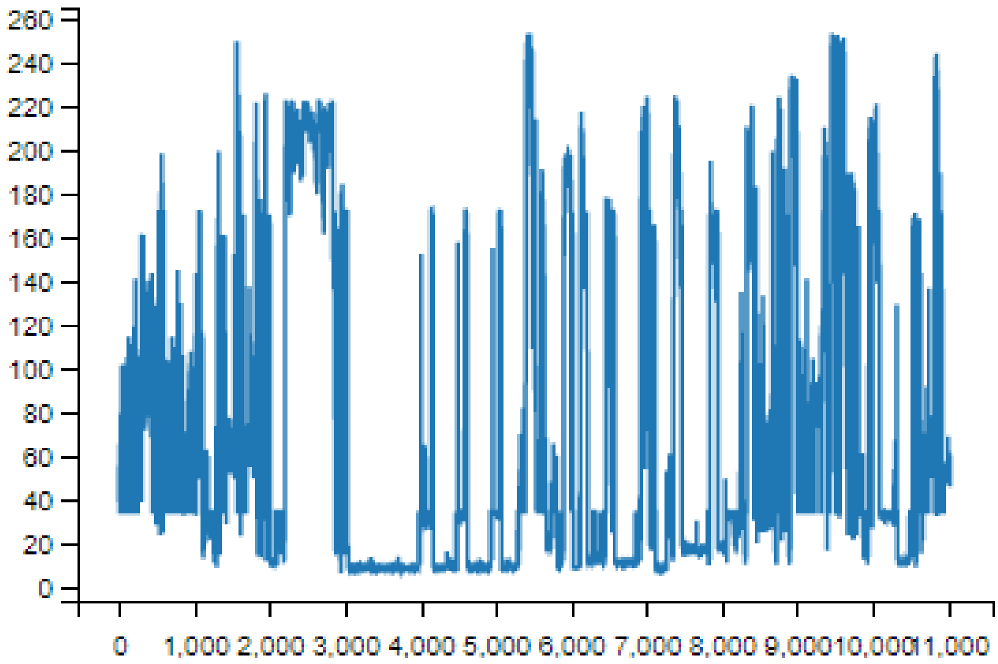

Figure 27 Predicted Power of target appliance

Plotting the predicted power against the testing target appliance power we get the following plot:

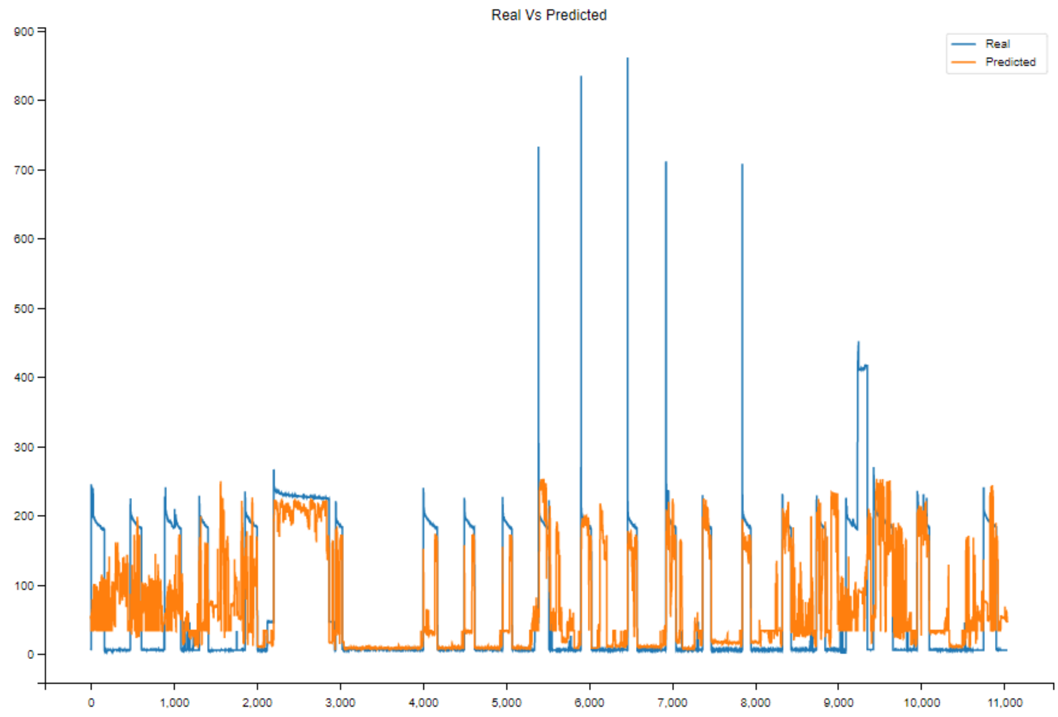

Figure 28 Predicted target power Vs test target power

In terms of metrics, the results obtained are shown in the following table:

True Positives	2933
False Positives	1690
True Negatives	5752
False Negatives	664
Recall	82%
Precision	64%
F1 Score	72%
Estimated Accuracy	68%

4.2.1.2 LSTM

After the training of the proposed LSTM model, the predicted power of the target appliance obtained can be visualized as follows:

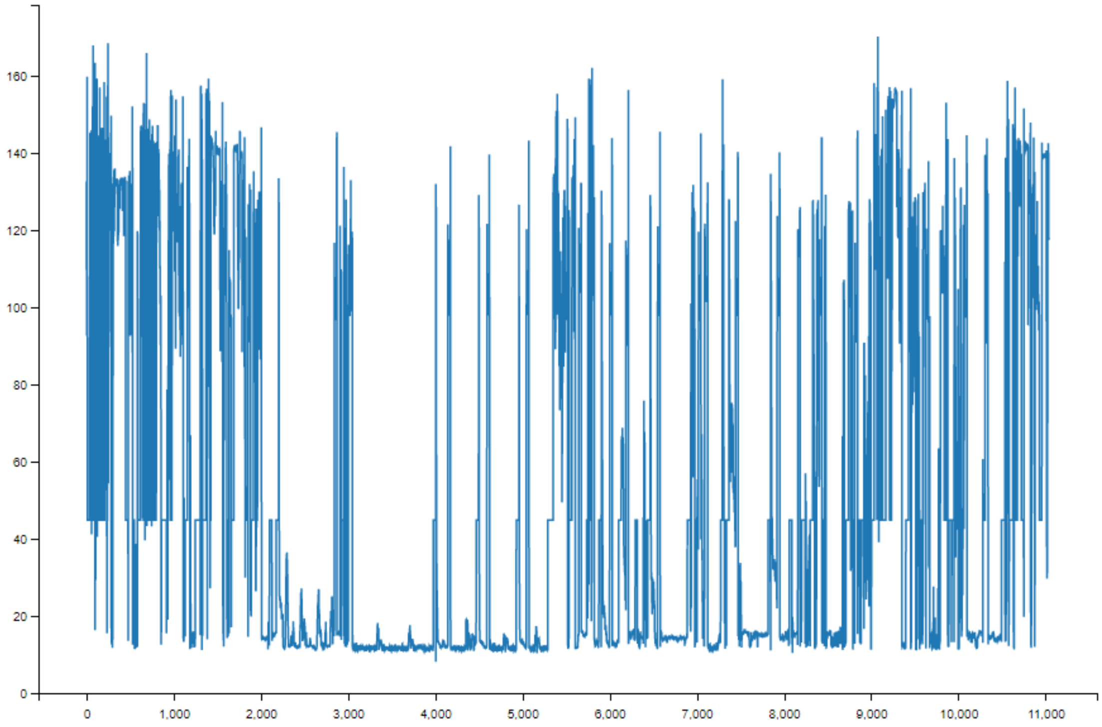

Figure 29 Predicted Target Power LSTM

Plotting the predicted power against the testing target appliance power we get the following plot:

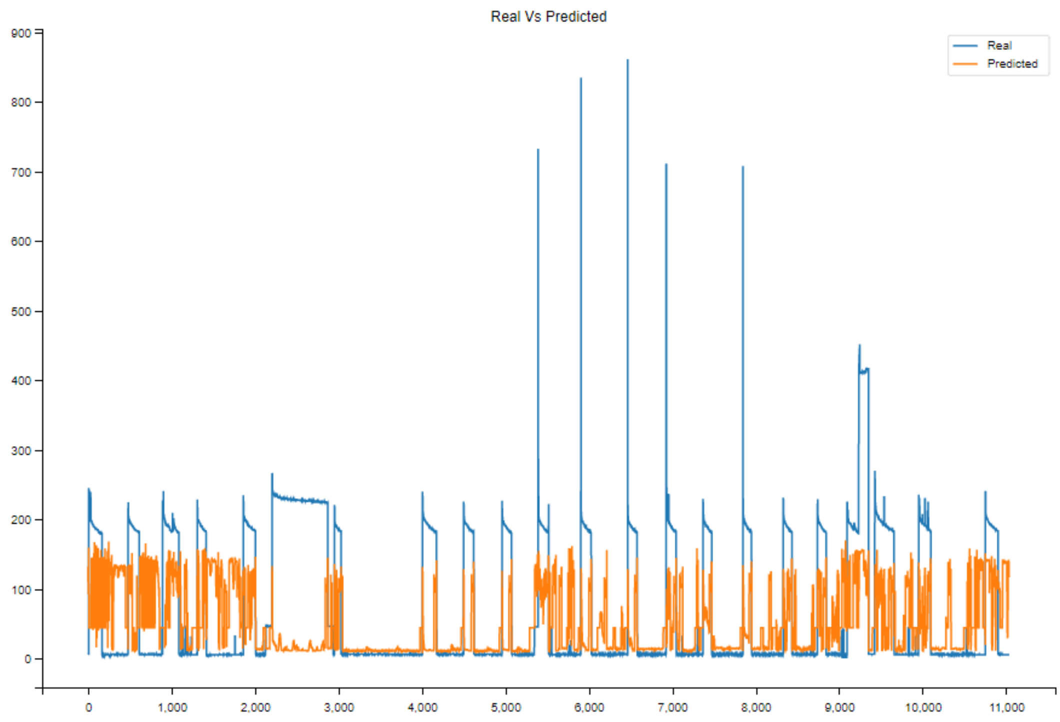

Figure 30 Predicted target power Vs test target power

In terms of metrics, the results obtained are as the following table describes:

True Positives	2980
False Positives	3777
True Negatives	3389
False Negatives	893
Recall	77%
Precision	51%
F1 Score	57%
Estimated Accuracy	53%

4.2.2 Conclusions

The results of the proposed CNN model outperformed the proposed LSTM model.

The results, carried out in the REDD dataset, in term of F1 score were 72% for the CNN and 51% for the LSTM. Furthermore, in terms of estimated accuracy, the results were 68% for CNN and 57% for the LSTM. In terms of True Positives, the CNN model had 1690 and the LSTM model had 2980.

The following table compares both proposed models CNN and LSTM:

	CNN	*LSTM*
True Positives	2933	2980
False Positives	1690	3777
True Negatives	5752	3389
False Negatives	664	893
Recall	82%	77%
Precision	64%	51%
F1 Score	72%	57%
Estimated Accuracy	68%	53%

5 Chapter 5: Applications using the NILM for IHEMS Project Dataset

5.1 Experimental Setup

5.1.1 Experimental Setup

The data used from the NILM for IHEMS dataset was from a private independent house in Algarve, South of Portugal and the target appliance was the Fridge. The sampling rate for both the aggregated power and the target power was 1 second. Below are some graphs visualizing both the aggregated power and the target appliance power.

Below is the aggregated power plot where the x-axis represents the power consumption in Kw and the y-axis represents time. The scale used is:

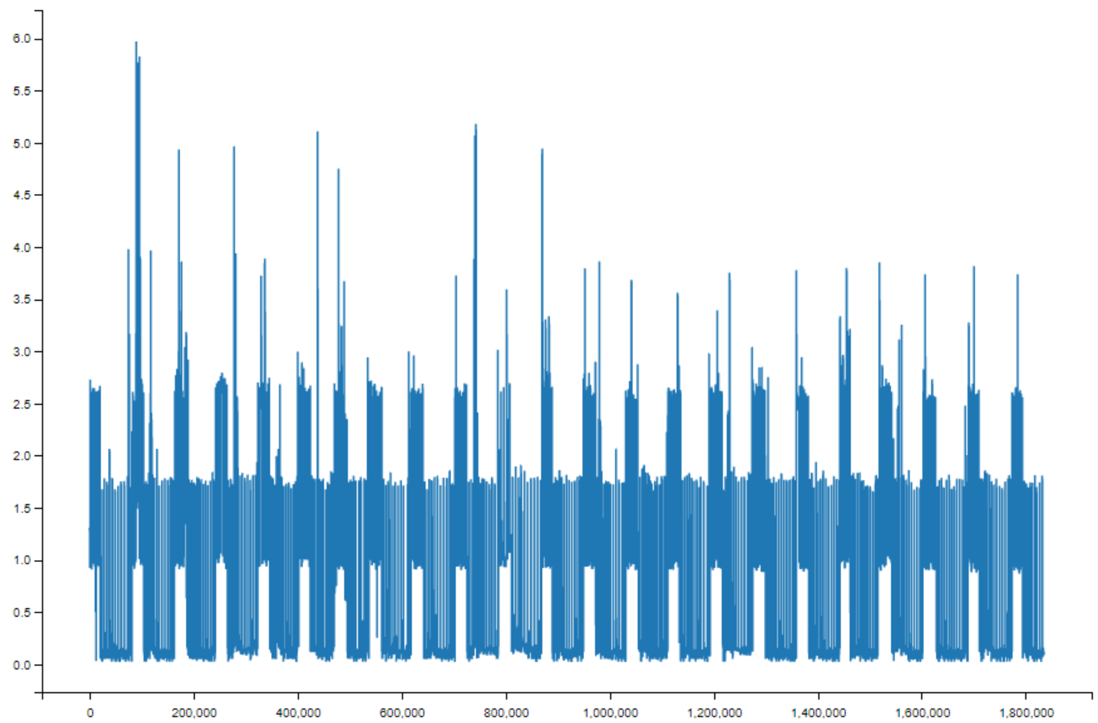

Figure 31 NILM for IHEMS Aggregated Power

A closer look into the aggregated power graph:

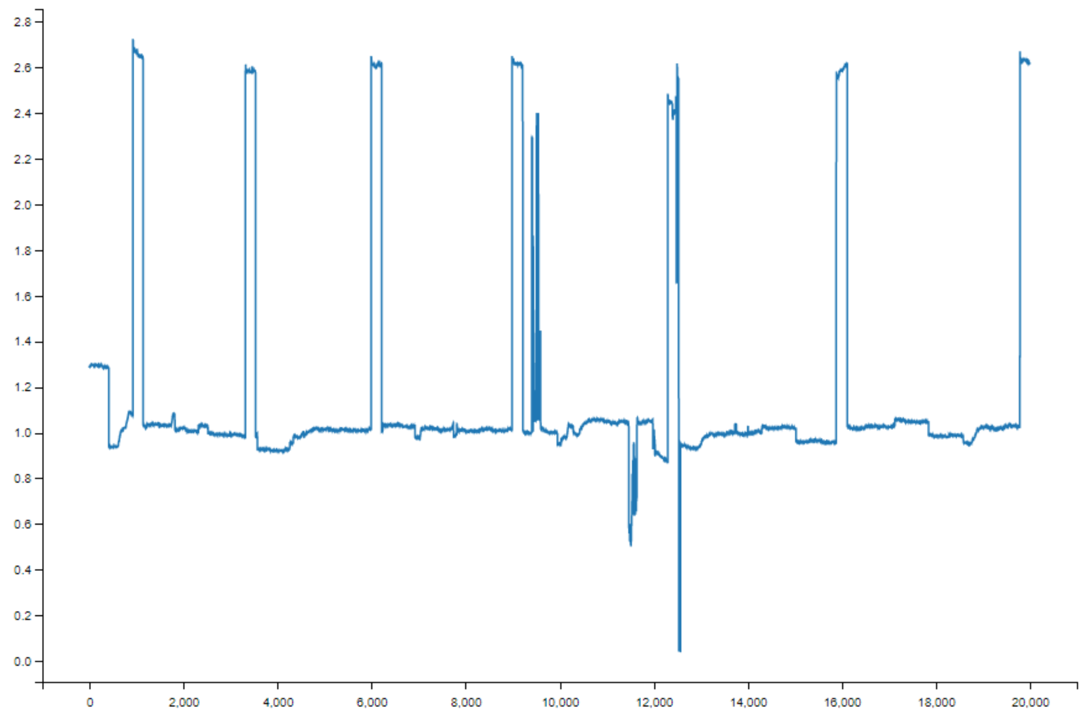

Figure 32 Zoom into the NILM for IHEMS aggregated power

Target Appliance Power in our case is the Fridge:

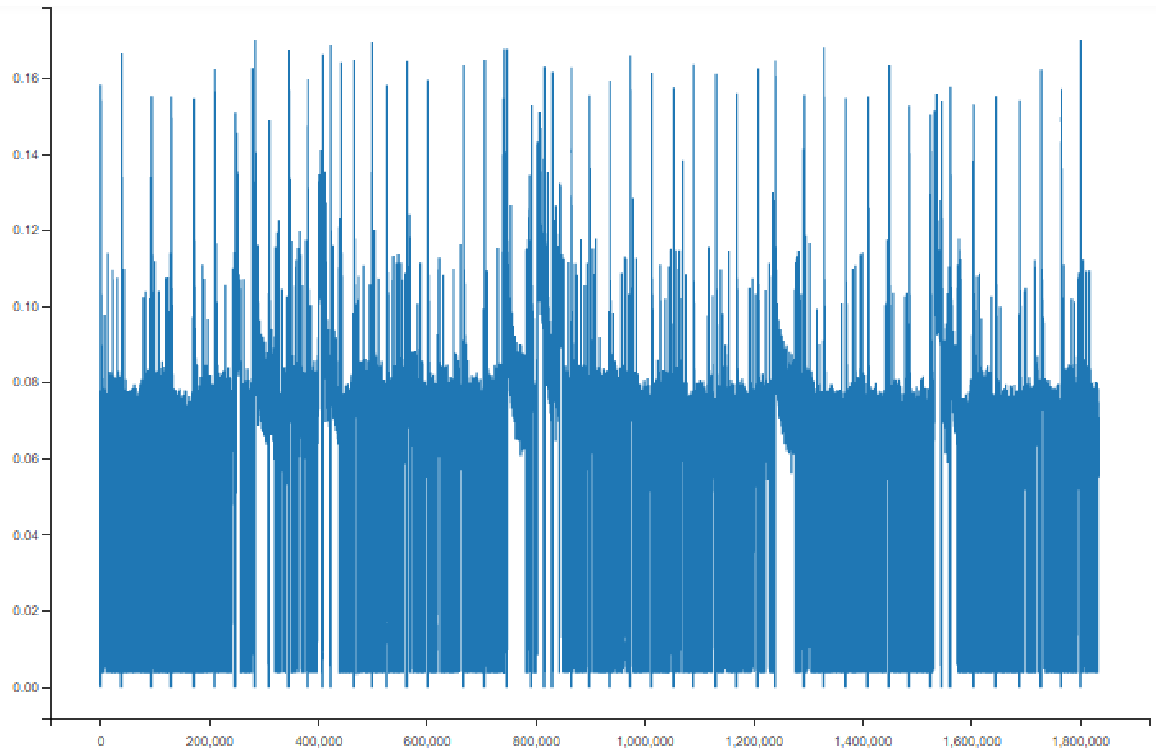

Figure 33 Target Appliance Power

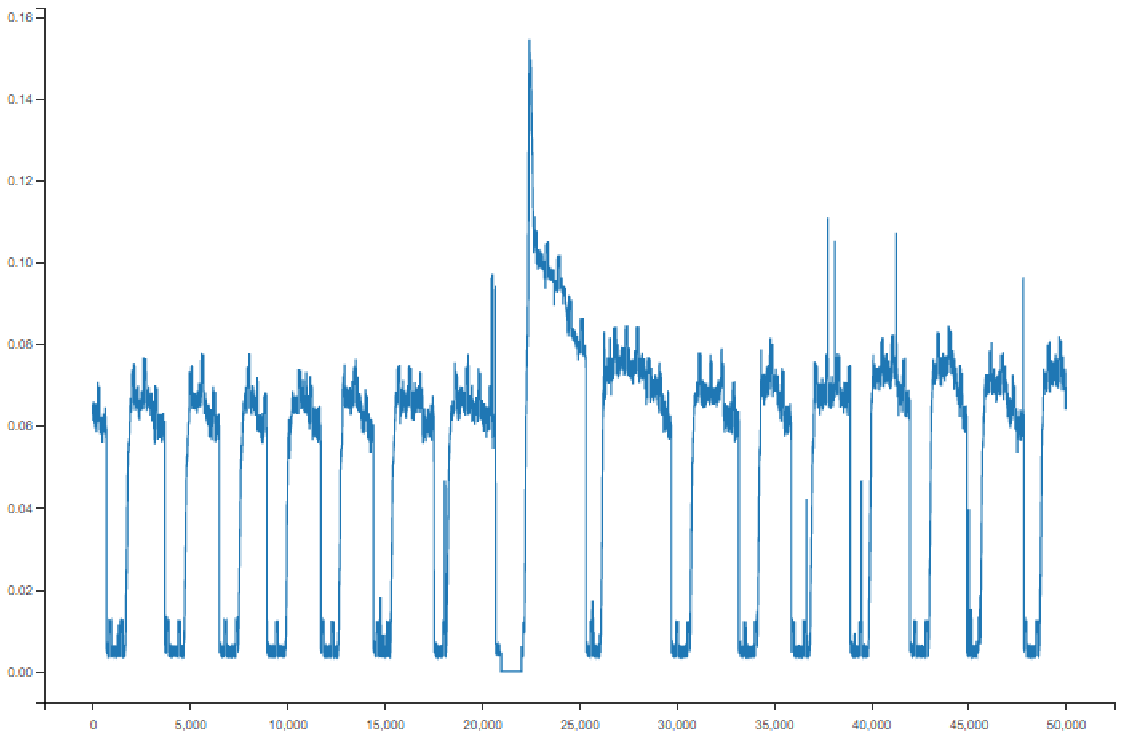

Figure 34 Zoom into the Target appliance power

The chosen dates windows were:

Dataset	*Start*	*End*
Training dataset	2021-06-02	2021-06-25
Testing dataset	2021-06-26	2021-06-30

The table is citing some parameters used in our experiment:

Parameter	*Value*
Sliding Window	30
Epoch	1000
Batch size	128
Early Stopping	Patience= 200

5.2 Results

5.2.1.1 CNN

After the training of the proposed CNN model, the predicted power of the target appliance obtained can be visualized as follows:

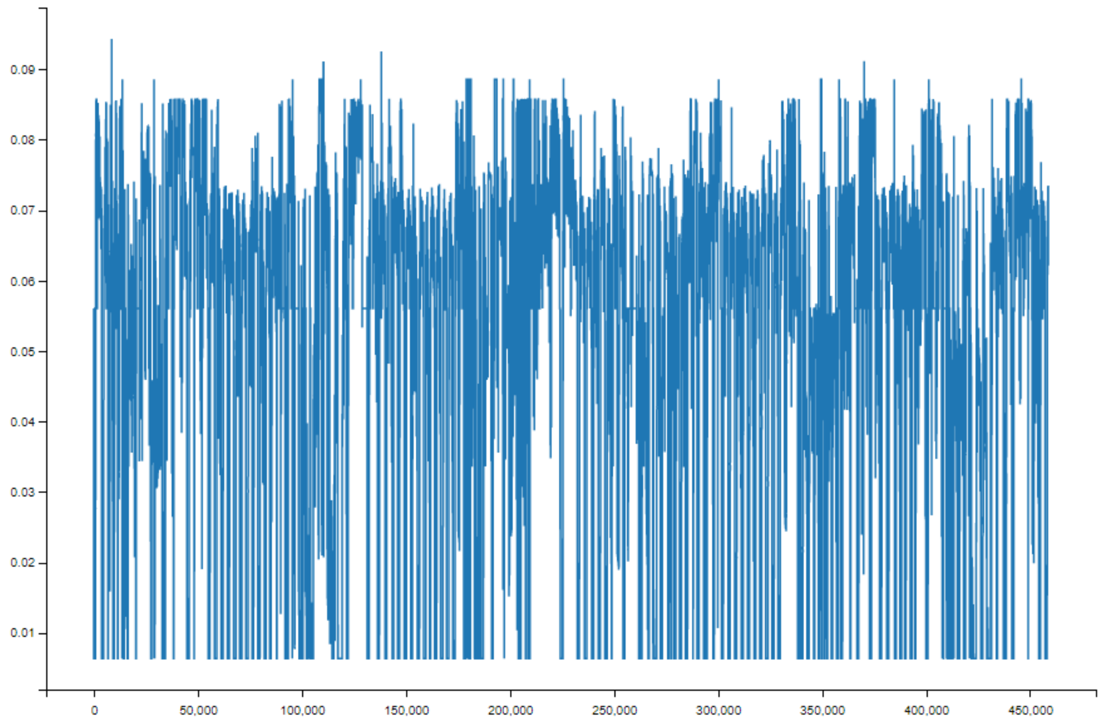

Figure 35 Predicted Power of target appliance

Plotting the predicted power against the testing target appliance power we get the following plot:

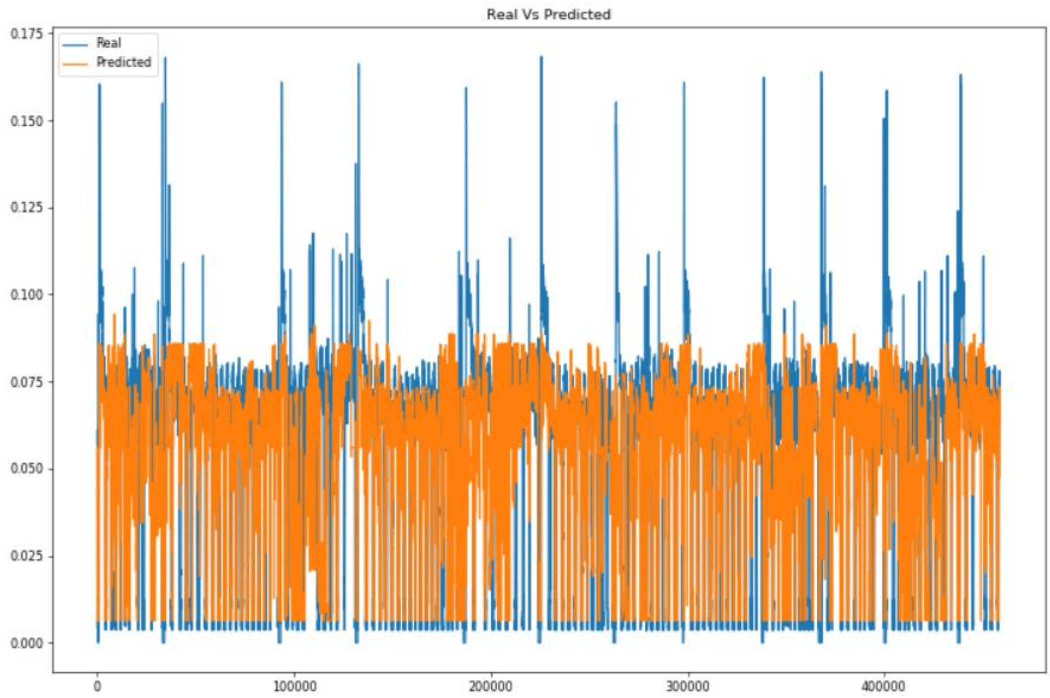

Figure 36 Predicted target power Vs test target power

If we zoom a little deeper we get:

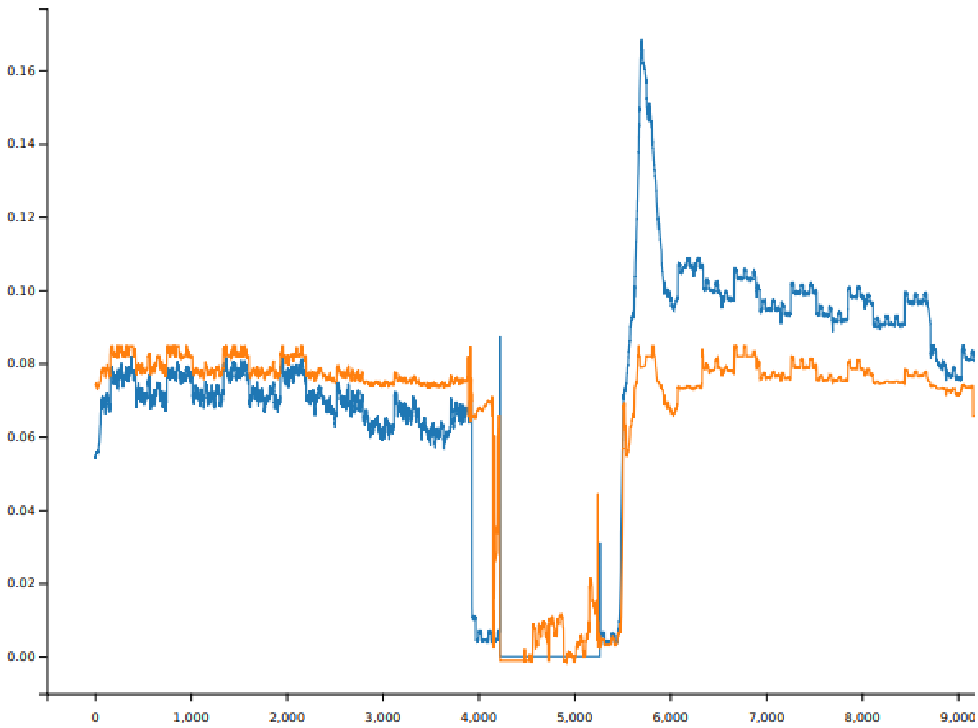

Figure 37 Zoomed Predicted target power Vs test target power

In terms of metrics, the results obtained are as the following table describes:

True Positives	304270
False Positives	37150
True Negatives	93739
False Negatives	23574
Recall	93%
Precision	90%
F1 Score	91%
Estimated Accuracy	87%

5.2.1.2 LSTM

After the training of the proposed LSTM model, the predicted power of the target appliance obtained can be visualized as follows:

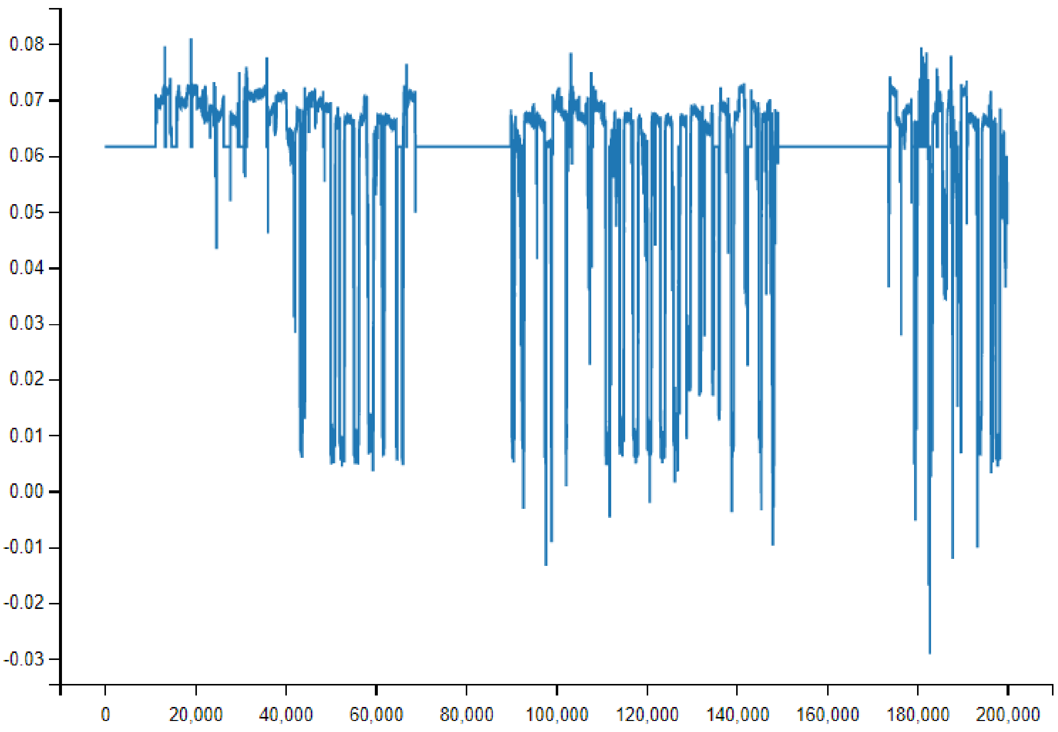

Figure 38 Predicted Target Power LSTM

Plotting the predicted power against the testing target appliance power we get the following plot:

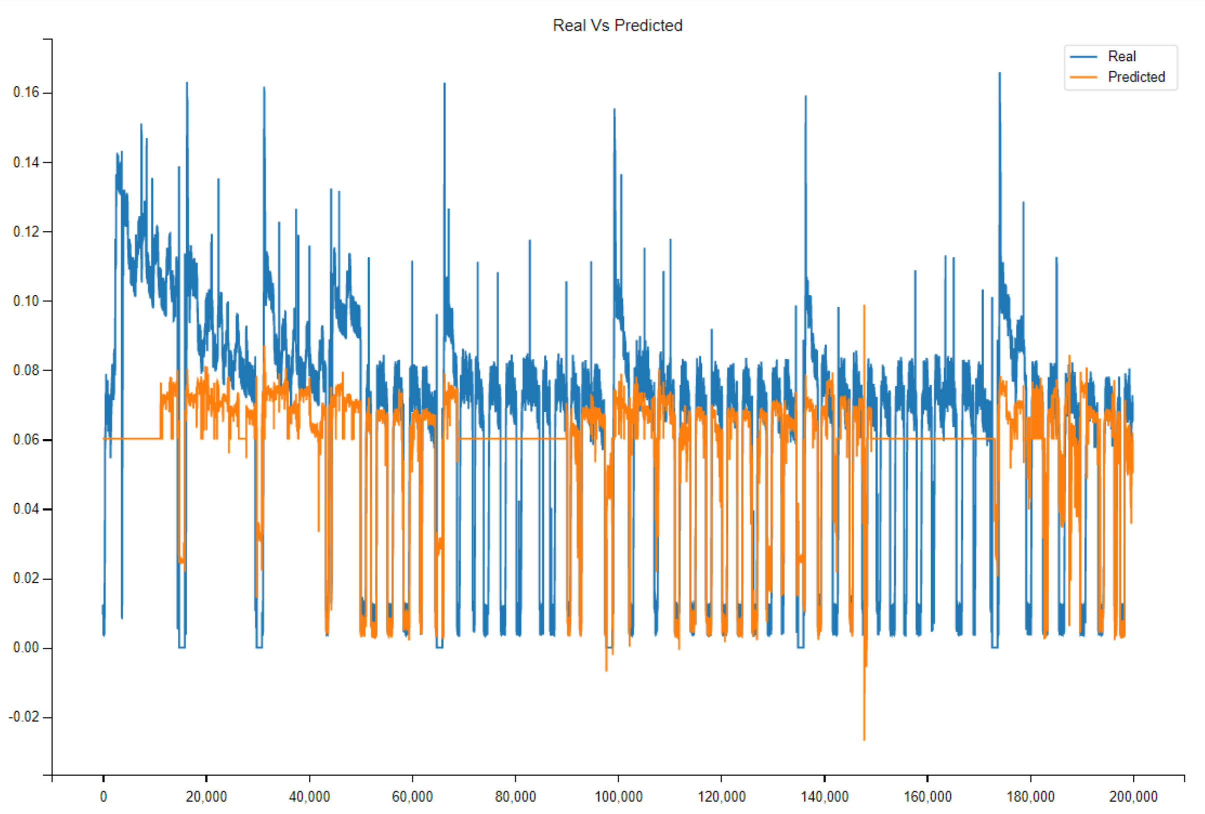

Figure 39 Predicted target power Vs test target power

If we zoom a little deeper we get:

Figure 40 Zoomed Predicted target power Vs test target power

In terms of metrics, the results obtained are as the following table describes:

True Positives	152847
False Positives	21092
True Negatives	25590
False Negatives	451
Recall	99%
Precision	88%
F1 Score	93,4%
Estimated Accuracy	86%

5.3 Conclusions

The results of the proposed LSTM model outperformed the proposed CNN model.
The results, carried out in the dataset recorded in the frame of the NILM for IHEMS project, in term of F1 score were 91% for the CNN and 93.4% for the LSTM. Furthermore, in terms of estimated accuracy, the results were 87% for CNN and 86% for the LSTM. In terms of True Positives, the CNN model had 37150 and the LSTM model had 152847.
The following table compares both proposed models CNN and LSTM:

	CNN	*LSTM*
True Positives	304270	152847
False Positives	37150	21092
True Negatives	93739	25590
False Negatives	23574	451
Recall	93%	99%
Precision	90%	88%
F1 Score	91%	93,4%
Estimated Accuracy	87%	86%

6 Chapter 6: Conclusion and Future Work

The work on this thesis was not easy with all the events that happened in our world in 2021. Unfortunately, the initial plans of this work were more in depth but they had to be changed along. This thesis is part of the project NILM for IHEMS. Its work focused on non-intrusive load monitoring using two deep learning technics. The models used are the following:

- Convolutional Neural Network
- Long Short-term Memory Recurrent Neural Network

Using two public datasets open for research:

- UKDale
- REDD

And the dataset collected in the private home subject to research in the project frame.

As future work, we are planning on improve and better tune the models, as well as to apply the models on further appliances.

7 Bibliography

A. Ruano, K. Bot, and M. Graça Ruano, (2020). Home Energy Management System in an Algarve Residence. First Results. Springer Nature Switzerland AG 2021 J. A. Gonçalves et al. (Eds.): CONTROLO 2020, LNEE 695, pp. 332–341, 2021.https://doi.org/10.1007/978-3-030-58653-9_32

Abdolmaged Alkhulaifi, Abdulah J. Aljohani (2020). Investigation of Deep Learning-based Techniques for Load Disaggregation, Low-Frequency Approach. International Journal of Advanced Computer Science and Applications, Vol. 11, No. 1, 2020

Alex Graves, et al., Framewise Phoneme Classification with Bidirectional LSTM and Other Neural Network Architectures, 2005.

Beckel C., Kleiminger W., Cicchetti R., Staake T., Santini S. 2014. The ECO Data Set and the Performance of Non-Intrusive Load Monitoring Algorithms. In Proceedings of the 1st ACM Conference on Embedded Systems for Energy-Efficient Buildings—BuildSys ’14, Memphis, TN, USA, 5–6 November 2014; pp. 80–89.

Gao, J., Giri, S., Kara, E. C., & Bergés, M. (2014). Plaid: A public dataset of high-resolution electrical appliance measurements for load identification research: Demo abstract. Proceedings of the 1st ACM conference on embedded systems for energy-efficient buildings, 198–199.

Gopinatha, R., Kumara, M., Prakash Chandra Joshuaa, C., & Srinivasa, K. (2020). Energy management using non-intrusive load monitoring techniques – state-of- the-art and future research directions. Elsevier Ltd. https://doi.org/10.1016/j.scs.2020.102411

G. W. Hart, "Nonintrusive appliance load monitoring," in Proceedings of the IEEE, vol. 80, no. 12, pp. 1870-1891, Dec. 1992, doi: 10.1109/5.192069.

Hasan Rafiq, Xiaohan Shi, Hengxu Zhang, Huimin Li, Manesh Kumar Ochani, (2020). A Deep Recurrent Neural Network for Non-Intrusive Load Monitoring Based on Multi-Feature Input Space and Post-Processing. Energies 2020, 13, 2195; doi:10.3390/en13092195. www.mdpi.com/journal/energies.

Huan Chen, Yue-Hsien Wang, Chun-Hung Fan (2020). A convolutional autoencoder-based approach with batch normalization for energy disaggregation. Springer Science+Business Media, LLC, part of Springer Nature 2020

J. Zico Kolter and Matthew J. Johnson. REDD: A public data set for energy disaggregation research. In proceedings of the SustKDD workshop on Data Mining Applications in Sustainability, 2011.

Kelly, J., & Knottenbelt, W. (2015b]). The UK-DALE dataset, domestic appliance-level electricity demand and whole-house demand from five UK homes. Scientific Data, 2, 150007.

Kolter, J. Z., & Johnson, M. J. (2011). REDD: A public data set for energy disaggregation research. Workshop on data mining applications in sustainability (SIGKDD), Vol. 25, No. Citeseer, 59–62.

Lawrence, S., Giles, C. L., Tsoi, A. C., and Back, A. D. (1997) "Face recognition: A convolutional neural-network approach." IEEE transactions on neural networks 8 (1): 98-113.

Luca Massidda, Marino Marrocu, Simone Manca (2020). Non-Intrusive Load Disaggregation by Convolutional Neural Network and Multilabel Classification. Appl. Sci. 2020, 10, 1454; doi:10.3390/app10041454. 21 February 2020

Michael Phi (2018). Illustrated Guide to LSTM's and GRU's: A step by step explanation. Towards data science. https://towardsdatascience.com/illustrated-guide-to-lstms-and-gru-s-a-step-by-step-explanation-44e9eb85bf21

Sepp. Hochreiter, Jurgen Schmidhuber (1997). Long Short-Term Memory Neural Computation 9(8):1735{1780, 1997

Shyamal Patel, Johanna Pingel. Introduction to Deep Learning: What Are Convolutional. Recorded March 2017 Neural Networks? https://www.mathworks.com/videos /introduction-to-deep-learning-what-are-convolutional-neural-networks--1489512765771.html

Weicong Kong, Zhao Yang Dong, Bo Wang, Junhua Zhao, Jie Huang (2020). A Practical Solution for Non-Intrusive Type II Load Monitoring Based on Deep Learning and Post-Processing. IEEE TRANSACTIONS ON SMART GRID, VOL. 11, NO. 1, JANUARY 2020