ASHESI

ASHESI UNIVERSITY

ELECTRICITY BILL CALCULATION FOR COMPOUND HOUSES IN GHANA USING ARTIFICIAL NEURAL NETWORK (ANN)

CAPSTONE PROJECT

B.Sc. Electrical and Electronics Engineering

Capstone Project submitted to the Department of Engineering, Ashesi University in partial fulfilment of the requirements for the award of Bachelor of Science degree in Electrical and Electronics Engineering.

Joel Anaafi Ampomah Nkansah

2021

ASHESI UNIVERSITY

ELECTRICITY BILL CALCULATION FOR COMPOUND HOUSES IN GHANA USING ARTIFICIAL NEURAL NETWORK (ANN)

CAPSTONE PROJECT

Capstone Project submitted to the Department of Engineering, Ashesi University in partial fulfilment of the requirements for the award of Bachelor of Science degree in Electrical and Electronics Engineering.

Joel Anaafi Ampomah Nkansah

2021

Table of Contents

DECLARATION

I hereby declare that this capstone is the result of my own original work and that no part of it has been presented for another degree in this university or elsewhere.

Candidate's Signature:

....

Candidate's Name:

....

Date:

....

I hereby declare that the preparation and presentation of this applied project were supervised in accordance with the guidelines on supervision of applied project laid down by Ashesi University.

Supervisor's Signature:

....

Supervisor's Name:

....

Date:

....

ACKNOWLEDGEMENT

First of all, I want to thank God Almighty for giving me the ability to write this paper. I am eternally grateful to Him. I would also like to thank my lovely parents who have supported me through this journey as well their consistent encouragement. This project would not have been possible without technical guidance from my supervisor, Francis Gatsi. My special thanks to Nana Akua Sereboo, Isaac Osei Nyantakyi and Jojoe Ainoo for the immense advice on data analysis and system implementation. Lastly, to everyone who has in one way or the other made this project a success, I appreciate you.

ABSTRACT

In Ghana, people live in a housing arrangement where members (tenants) of the household share basic facilities such as washrooms, kitchens and electricity meter this is called a compound house. Over the years, there have been several complaints regarding these shared facilities but one remains paramount and is the electric meters. Tenants mostly have concerns about how electricity bills are shared amongst themselves. It is assumed that tenants may not consume electric power as much as others do and as a result conclude bills are not shared fairly.

Previous research has been done to verify the ability to split bills by implementing the Non-Intrusive Load monitoring system with focus on identifying different appliances used by different tenants. However, this is not take into consideration similar appliances used by different tenants.

This paper investigates the possibility of distinguishing between similar appliances used by different tenants. It was discovered that the Time Series Classification Algorithm had a higher overall performance when compared to the Multilayer Perceptron (MLP) hence is efficient for distinguishing between similar appliances used by tenants.

Chapter 1: INTRODUCTION

1.1 Introduction/Background

Energy is the basic necessity for the economic development of a country. In modern times, the availability of a huge amount of energy has birthed higher agricultural and industrial production, better transport facilities, etc. The greater the per capita consumption of energy in a country, the higher the standard of living of its people. Energy exists in various forms in nature but the most important form of it all is electrical energy[1][2].

Electrical energy is a very convenient form of energy as it can be easily converted into other forms of energy. It is much cheaper than other forms of energy and it is not associated with smoke, fumes or poisonous gases ensuring its cleanliness.

1.2 Electricity in homes (The Ghanaian Home)

Every single household requires the supply of electricity to be able to go about their daily activities. These activities may include washing, where a device such as a washing machine may be used. Some also deploy the use of other electrical devices such as electrical stoves, refrigerators, TV sets, etc. which require the supply of electricity.

Ghana generates electric power from different sources such as hydropower, thermal energy(fossil-fuel) and some renewable energy sources. Electricity is generated at various power plants and then is transmitted at high voltages to substations where the power is stepped down to a nominal voltage and then distributed to various homes. Electricity distribution and sale services are currently conducted by three companies: Electricity Company of Ghana (E.C.G) which is in charge of delivering power to people in the southern part of Ghana, Northern Electricity Distribution Company (NEDCo) which provides the Norther parts of

Ghana with electricity and Enclave Power Company Ltd, a privately owned electricity distribution company that delivers electricity within the Free Zones Enclave at Tema Industrial Area[3].

The provision of electricity to homes all around the nation does not come free and therefore needs to be paid for. The utility providers, Electricity Company of Ghana provide consumers with meters to enable the billing process according the amount of energy or power consumed. These energy meters are calibrated in billing units (kilowatt hour [kWh]). Meters of different accuracy classes are used for different purposes and applications (e.g. Residential, Non-residential, Industrial, etc.) based on the accuracy of the requirement[4].

This type of metering system poses a challenge for residential arrangements, specifically compound houses. A compound house in a Ghanaian setting is a human habitat referring to a cluster of buildings in enclosure, having shared utilities and amenities such as water, bathrooms, kitchens, electric meters, etc.[5] When it comes to the payment of some these utility bills, a lot of confusion arises as a result of the difficulty in sharing the amount each person pays.[6]

The distribution of electricity to compound houses faces this same situation as members in compound houses (popularly known as tenants) find it difficult to share electricity bills because of their shared electric meter[7]. Land-lords (owners of compound houses) usually share the bills based on the number and types of electrical equipment one has in his or her household. This results in heated arguments and quarrels as people believe others would have consumed more power.

To resolve this problem of having to share an electric meter, some tenants arrange with land-lords and the power distribution service for instance E.C.G to obtain separate meters for their houses[7]. This method of separation is costly as it would involve rewiring of the house as well as the purchase of a new meter.

1.3 Problem Definition

Compound houses face the issue of a shared meter and therefore tenants find it difficult to agree with the electricity bill each of them is required to pay. These compound houses have their rooms being wired together (the electricity network) and therefore make it difficult to differentiate who consumes the most amount of power. Land-lords are therefore forced to share bills based on the electrical appliances and equipment used by tenants in the house. However there exist a greater setback when different tenants may use similar appliances (same make and model) and this paper seeks to address this issue.

1.4 Objectives of the Project Work

This project seeks to:

a. Investigate the possibility of distinguishing similar appliances used by different tenants in a compound house using an appropriate technique or algorithm
b. Distinguish similar electrical signatures of appliances

1.5 Motivation for Project Topic

The use of shared meters by several compound houses is one issue that I have always want to address. This because of how unfairly the bills of tenants are calculated unfairly. I have always wanted to research in this area of metering systems as I have witnessed the confusion it creates among tenants in a compound house. The Electricity company of Ghana as of now has not found a way to address that issue and from my short experience working with them, it does not look like something would be done about it soon. This setback has been in existence for a while and till date people still have several complaints about using a shared meter and would prefer having separate meters however that is costly.

1.6 Requirements

This project would focus on two main requirements. These are the user requirements and system requirements.

1.6.1 User Requirements

To be able to undertake this investigation on how similar appliances used by different tenants could be distinguished, it is necessary to undertake an extensive research to confirm from users whether this case is actually a problem. It is important to be there with the people to understand their concerns.

In January and February 2021, I conducted interviews with tenants in 8 different compound houses at Dome and Madina respectively. These interviews were held to help me gather information on the problem. The Electricity Company of Ghana are the main suppliers of electricity to these compound houses. Some of these houses have been provided with pre-paid meters and other post-paid meters. During the interview at Dome, one tenant mentioned that he shared the pre-paid meter with other tenants but he claimed he uses few appliances. For a pre-paid meter an amount is paid and the credit is issued onto the system before the meter works. Based on the consumption of power by those sharing the meter, the credit can get exhausted and these are the issues some tenants are facing. Similar to the case of this tenant whose credit gets exhausted within a short period. The only prediction, the tenant could make is that the other tenants use more appliances.

The interviews conducted with some tenants at Madina also shared similar sentiments. In this case, people were resistant to pay a contributing price since they do not consume the same of amount of power. As a result of all this confusion, some tenants decided to purchase their own meters and this is costly. It is almost impossible for all people in similar situations to purchase separate meters.

The interviews conducted has shown that electricity bill calculation is still a problem for compound houses in Gama today and as such requires a system to mitigate this problem.

1.6.2 System Requirements

For the proposed approach, the deliverable should:

- be able to distinguish between similar appliances used by different tenants
- be able to detect on/off state of appliances
- be able to differentiate the timing patterns between similar appliances
- use a non-intrusive method to collect data from different appliances from the main utility entry point. It should accurate receive current drawn by different appliances in real time

1.7 Proposed Solution/Scope of Work

This project investigates the possibility of being able to differentiate similar appliances used by different tenants as this could also help in achieving a more accurate bill calculation for tenants in compound houses.

The scope of this project looks at proposing appropriate techniques and algorithms that would be able to help distinguish between similar appliances used by the different tenants.

Chapter 2: Literature Review

In this chapter, there would be discussions on related work and some scientific papers in this specific area of electricity billing calculation, non-intrusive load monitoring and differentiating between similar appliances. This would help provide information on how the problem of electricity billing calculation for compound houses in Ghana can be solved leading to a more accurate solution.

Compound houses in Ghana have several tenants share important utilities and one example is an electric meter and this would be our focus[4]. These houses mostly have complex wiring. [6] looked at how compound houses that share one meter could have a separate wirings and how the consumption of each houses could be measured and calculated. He therefore built a smart energy monitor with the assumption that the wiring of rooms was separated. [6] however mentioned that his assumption is not usually the case as most compound houses have joint wiring and therefore makes it difficult to measure the amount of power consumed by tenants in a compound house.

To solve this, a concept known as Non-Intrusive Load Monitoring (NILM) is introduced. NILM is used to estimate the consumption of individual household appliances based on the aggregate consumption of a home by analyzing voltage and current that goes into a house[8]. This system disaggregates total electrical usage into appliance related signals. This is done by using Machine learning techniques. These machine learning techniques can be divided into two categories and these are supervised and unsupervised learning algorithms[9]. NILM usually adopts the supervised learning algorithm[10]. This type of algorithm requires a large number of labeled data when training the system. However, the unsupervised learning algorithms are also deployed in NILM. This approach does not require individual appliance data and the models information is captured without any human intervention[11]. NILM is made up of four major stages and these are; Data Collection, Event detection, Feature Extraction and Load Identification[12].

Collecting data is important for testing algorithm and comparing performance of various results obtained[9]. With the help of smart meters and some sensors data can be collected. These sensors measure parameters such as real power, reactive power, apparent power, power factor and even the frequency. The event detection stage of NILM analyzes how these parameters change over time by assessing their steady and transient state characteristics. Based on the type of appliance, event detection may vary. [13] identified four appliances types and there are; simple on/off, finite state machine (FSM), constantly on and continuously variable. Simple on/off are the easiest to detect and they are only ever on or off. The different features are gained from the results in the event detection stage. It is at the feature extraction that we obtain the parameters necessary for load identification. During the load identification stage, machine learning techniques are used in disaggregation. Others also make use of Signal Processing techniques but the machine learning algorithms are mostly used.

In the Ghanaian context the implementation of NILM has not been much as this is an area in which more research is being done. [10] aims at verifying if the Feed Forward Artificial Neural network (FANN) can be used to split bills by implementing a Non-Intrusive Load Monitoring System for compound Houses in Ghana. The dataset used by [10] provides only apparent power as a parameter for the load identification algorithm. The different appliances are clustered per tenant and weighted based on their apparent power consumption[10]. The apparent power consumption for appliances that are on per tenant is calculated by finding the ratio of consumption. This proposition by [10] shows how different appliances used by different tenants can be identified hence their power consumption calculated.

There is a huge possibility that different tenants may use similar appliances. This can cause problems when trying to provide information on the power consumption of appliances for the different tenants. In an attempt to solve this, a poorly designed energy-disaggregation algorithm might erroneously cluster identical devices into a single group. Notwithstanding, appliances with the same make and model have some properties that enable distinguishing to be possible[14]. The timing patterns and electrical signatures of identical appliances could be used to distinguish them[12].

Similar appliances used by different tenants may not necessarily have an on or off time as tenants may use these appliances at different times during the day. Distinguishing features such as wattage-based observations and cycling time is used by [12] to analyze the timing patterns of the appliances. Although similar appliances might have similar electrical signatures, it is highly impossible that these two or three similar appliances will cycle synchronously [12]. The physical properties of certain components (for e.g.; resistors or capacitors) used by appliances may vary and as a result reactive power and other parameters such as frequency would differ [12].

Other techniques may be used in identifying the differences between similar devices used by tenants. [12] provides a solution which automatically detects and classifies the use of devices in a home from a single point of sensing. Most of the devices assessed employ the use of switch mode power supplies (SMPS) and this generates continuously high frequency electromagnetic interference (EMI) throughout the wiring[15]. Similar devices would produce identical EMI but may look different since the devices are attached at different points along the power line[13]. The spectrums of the similar appliances are then observed and their amplitudes are used to differentiate between them.

Chapter 3: Methodology

This project investigates the possibility of being able to differentiate between two similar appliances used by different tenants in a compound. Based on research, there are different ways by which similar appliances could be distinguished. Common parameters could be used to differentiate electrically similar devices and this includes; current harmonics and transient properties such as transient response time, current spikes and spectral envelopes [16]. Also, [13], mentions that electrical signatures could be used as clues to detect any two similar or identical appliances.

This project would focus mainly on using timing patterns to differentiate similar appliances. By observing the rate at which power consumed and time at which it is done, timing patterns can be obtained. Although these similar appliances are said to be of the same make and model, it is highly unlikely that the appliances of two different tenants will cycle synchronously (i.e. turn on and off at the exact same time) [14].

In the graphs below, we see the power consumption of three different appliances namely; refrigerator, light and microwave in two different houses being assumed to be tenants and this is plotted against time. This shows the timing patterns of both houses. The data used in plotting this graph was obtained from the Reference Energy Disaggregation Data Set (REDD) and this data is being used in the project as well. The different time patterns shows how the appliances of the two different tenants do not cycle synchronously.

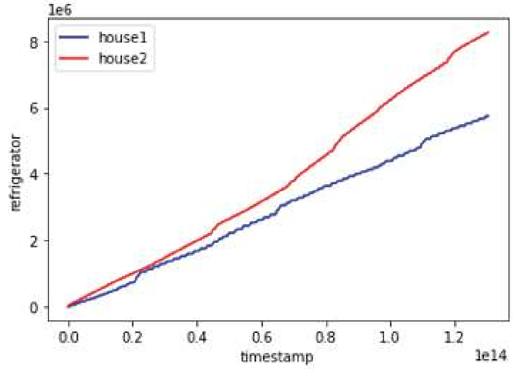

Figure 3.1 Graph of refrigerator

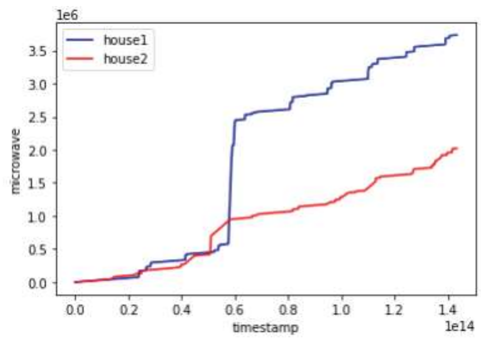

Figure 3. 2 Graph of Light

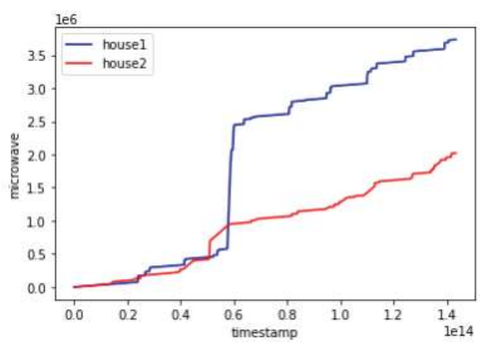

Figure 3.3 Graph of Microwave

The requirement of this project is to propose an approach to have an NILM system which is capable of differentiating similar appliances used by different tenants in a compound. This project would use a software-hardware approach as required by the NILM system.

3.1 Hardware Approach

The Hardware system involves designing a low-cost energy monitoring system that would measure the current of the different appliances. With the help of a non-invasive current transformer, current can easily be measure and sent to a microcontroller. This method allows the collection of local data. However, enough data could not be collected, hence referred to an online dataset called the Reference Energy Disaggregation Dataset (REDD). REDD Dataset is a freely available data set containing detailed power usage information from several homes which is aimed at furthering research on energy disaggregation. In this Data set there are about 6 houses and in each house contains several appliances that have been labelled. The apparent power of these appliances were measured. For this project House 1 and 2 is selected and the power consumption of the appliances; light, microwave and refrigerator is used as data.

3.2 Software Approach

The Software system proposes the use of two classification models under Artificial Neural Network. These two classification models allows the input data, in this case timing patterns (i.e. time stamps) and the respective power consumptions of the appliances to be classified as a predefined labelled classes. The two algorithms proposed are; the Time Series Classification algorithm (TSC) and the Multilayer Perceptron (MLP). For this project, the two algorithms are compared to each other and their performance is measured based on which algorithm is most appropriate for distinguishing between similar appliances used by different tenants.

3.2.1 Algorithms

1. Time Series Classification (TSC)

A time series classification model classifies data points over time based on its features. Essentially, this model helps to identify patterns within time series associated with relevant classes. A Sequential Neural Network mode for Keras (open source software library) is used to extract features on the data and find which houses the electric consumption of the similar appliances belongs to.

2. Multilayer Perceptron (MLP)

A Multilayer Perceptron model, MLP, is a standard fully connected neural network model. It comprises of layers of nodes where each node is connected to all the outputs from the previous layer and the output of each node is connected to all inputs for nodes in the next layer. The figure below shows a sample MLP neural network.

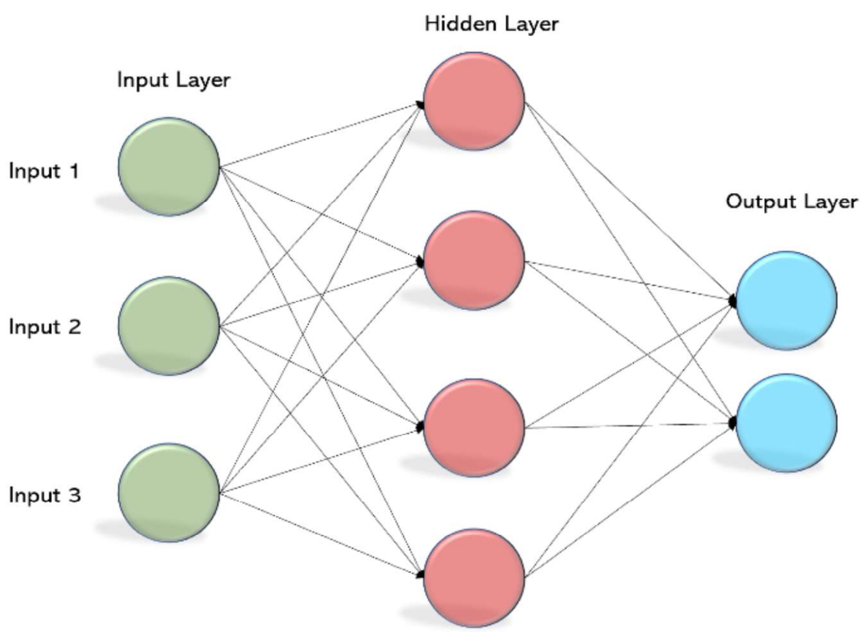

Figure 3.4 A sample MLP neural network

The MLP network is created with one or more dense layers. It works well with tabular data such as the REDD dataset, with one column for each variable.

3.2.2 Measuring performance

To measure the performance of the Time Series Classification algorithm (TSC) and Multiple Perceptron algorithm (MLP), the confusion matrix proposed by (cite) is used. A confusion matrix is a table used to describe the performance of a classification model on a set of test data for which the true values are known.

Chapter 4: Design and Implementation

This chapter reveals the experimental set-up used in implementing the NILM energy monitor and the appropriate Artificial Neural Network algorithm to split bills for different tenants who use the same appliances. An in-depth analysis on design decisions would be discussed in this chapter.

4.1 Assumptions

There are some assumptions that are being made to justify the implementation of software and hardware designs;

I. All appliances are type 1. This means the appliances are either off or on.

II. The compound house under focus is single phase.

III. The different tenants do not put their appliances on at the same time.

4.2 Hardware Design and Implementation

The entire makeup of this hardware design and the selection process of the different components is mainly their compatibility with Arduino, availability, cost and ease of use. A low cost NILM energy monitoring system would be built to be compatible with the software implementation. The selected components of this energy monitoring system are the ATMEGA 2560 microcontroller, a real time clock, a Non-invasive current Transformer sensor, an LM7805 voltage regulator, an LCD and a 9V rechargeable battery.

4.2.1 ATMEGA 2560 microcontroller

The Arduino Mega 2560 is a microcontroller board based on the ATMEGA 2560. This microcontroller was chosen based on its compatibility with the DS3231 real time clock and the current transformer sensor. It is also a powerful microcontroller that can fully support the machine learning algorithm.

4.2.2 Real time clock

It would keep an updated track of the current time since this project deals with timing patterns. The most available and compatible real time clock with the ATMEGA 2560 is the DS3231 real time clock.

4.2.3 Non-invasive current transformer sensor

This sensor is used to measure alternating current. It is a crucial component in the energy monitoring system as all electrical signals from the electric meter would pass through this component before it gets to the microcontroller for further computation. It is reduces the high level of current received to a smaller current so that it can be measured safely. This sensor receives an input current up to 100A and outputs between 0 to 50mA. It can be interfaced with Arduino and is compatible with the ATMEGA 2560, a major reason for its selection.

4.2.4 LM7805 voltage regulator

This is a voltage regulator that outputs +5V. This is used to step down the voltage that comes from the main power source so that it is compatible with the microcontroller since its input voltage is 5V.

4.2.5 Liquid Crystal Digital (LCD)

This would be used to display the power consumption of each tenant.

4.2.6 9V rechargeable battery

This would be the main power source for the energy monitoring system. This system should still work in case of any power outages so it is not dependent on the mains coming from the house.

4.2.7 Hardware Circuitry

The hardware circuitry of this proposed implementation can be seen in Figure 4.4. The overall circuitry involves the connection of a 9V rechargeable battery to power the circuit, an

ATMEGA 2560 which is used to execute the various instructions in Arduino, a real time clock and an LCD for display.

In the Power Circuit in Figure 4.4, a 9V battery is connected in parallel to a 10uF which smoothen the signals. It is then connected to the input of the LM7805 step-down voltage converter. At the output of the LM7805 converter, 5V is obtained.

The ATMEGA 2560 is boot-loaded to enable the use of the Arduino IDE to send instructions (code) to the microcontroller. A DS3231 real time clock is connected to the microcontroller. This component keeps track of time even if the microcontroller is reprogramed or disconnected from the main power source. The real time clock module can work on 5V and uses the I2C communication protocol which makes the connection to Arduino easy.

The non-invasive current transformer in this circuit measures the alternating current of different electrical loads. A burden resistor must be used in conjunction with the current transformer as it affects the output voltage, signal to noise ratio and the bandwidth of the signal[17]. Once the burden resistor increases, the output voltage value also increases. In Figure 4.4, a 33 Ω is placed parallel to the current transformer. The SCT-013-000 Non-Invasive AC Current transformer sensor can measure between 0 to 100 A AC, gives an output current between 0 to 50mA and has a turns ratio of 100A:0.05A [18]. The voltage across the burden resistor is 2.5V. The reference voltage from the microcontroller (5V) is potentially divided using two 10k Ω resistors in series. The mid-point of this connection is then connected to the transformer hence 2.5V. To obtain the suitable burden resistor, the mid-point voltage is divided by the maximum current the transformer can measure in its root mean square value (Irms). Hence the selection of the burden resistor with rating 33 Ω.

Table 4.1 DS3231 real time clock (RTC) connection to ATMEGA 2560

Pin on RTC	Pin on ATMEGA 2560	Description
VCC.1	5V	Power Supply
GND.1	GND	Ground
SDA.1	(SDA/INT1)PD1	Data lines
SCL.1	(SCL/INT0)PD0	Clock lines

Table 4.2 LCD 16x2 I2C connection to ATMEGA

Pin on LCD	Pin on ATMEGA 2560	Description
VCC	5V	Power Supply
GND	GND	Ground
SDA	(SDA/INT1)PD1	Data lines
SCL	(SCL/INT0)PD0	Clock lines

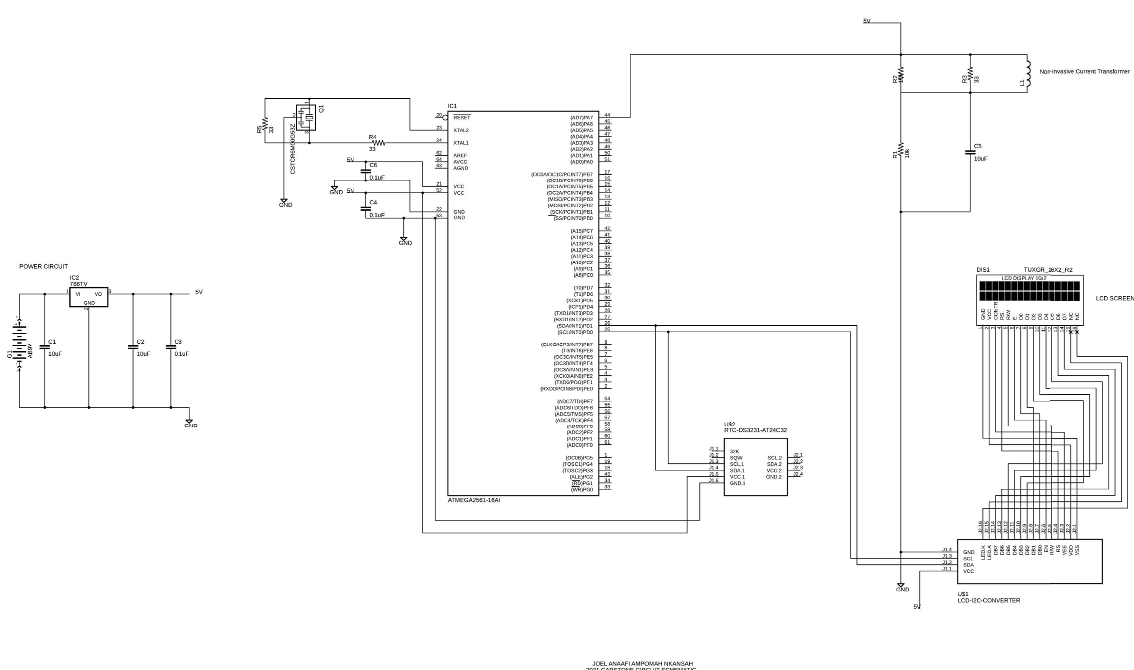

Figure 4.0.1Schematic of the hardware design of an NILM energy monitor

4.3 Software Design and Implementation

The proposed software design is an implementation of a classification model Artificial Neural Network (ANN) system to help identify the different timing patterns and further calculating the bills of the different tenants based on the power consumption of their appliances. The reason for the use of Artificial Neural Network is that ANNs learn to perform task by learning from examples. They learn from processing examples which usually contain an input and an output. This project uses ANN to learn from processing examples with timing patterns as inputs and tenants as outputs. The idea is to allow an ANN identify what new timing pattern belongs to a particular tenant based on trained processing examples. The flow diagram of the software design is shown in Figure 4.

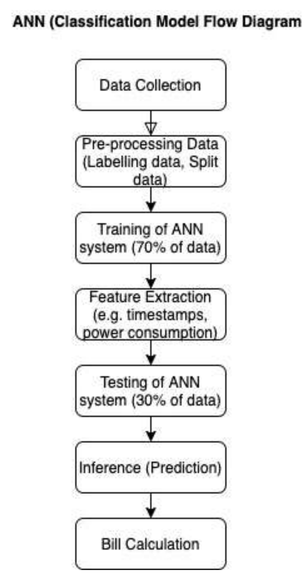

Figure 4.0.2 A Flow Diagram of the Artificial Neural Network model

4.3.1 Data Collection

In this project, data was used from an online data set called the Reference Energy Disaggregation Data Set (REDD). This is because of my inability to obtain data from our main electricity service providers as well as my inability to build a full hardware implementation for collecting my data from various compound houses. The data used in this project was collected from house 1 and house 2 in the REDD dataset. These two houses are assumed to be tenant A and tenant B. Three different appliances of which are assumed to be similar in terms of its make and model. These devices are; refrigerator, lighting and microwave.

4.3.2 Pre-processing Data

Data for two houses (house 1 and house 2) was used from the REDD dataset. For each house, the timestamp and the data on appliances including refrigerator, light and microwave were selected to be used in training the models. The output column had 2 classes: house 1 and house 2. The 2s in the output column were replaced with 0s so that the output column can be compared to the predicted values from the deep learning models.

4.3.3 Training and Testing of dataset

80% of the data is trained over both algorithms and 20% of the data is used for testing

4.3.4 Feature Extraction

During the training of datasets, several features such as timing patterns and power consumption of appliances is extracted for further analysis.

4.3.5 Prediction

During testing the system predicts which timing patterns belongs to which tenant as well as the appliances. This helps to differentiate the similar appliances used by the two tenants.

Chapter 5: Results and Discussion

This chapter outlines the method used for measuring performance of the algorithms implemented and how appropriate it is to distinguish between similar appliances used by different tenants. After building a machine learning model, it is important to determine how well the model works. The confusion matrix model would be used measure the performance of the ANN classification models.

Under the confusion matrix, an evaluation is done on the accuracy, precision, recall and F1 score of both algorithms. This is calculated using the formula below.

Table 5.1

	Positive Predicted	Negative Predicted
Actual Positive	True Positive (TP)	False Negative (FN)
Actual Negative	False Positive (FP)	True Negative (TN)

Formulas:

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

Precision is the percentage of predicted positives that were correctly classified.

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

Recall is the percentage of actual positives that were correctly classified

$$F1\ score = \frac{2\ \times Precision\ \times Recall}{Precision\ + Recall}$$

F1 Score is the harmonic mean of Precision and Recall. It gives a much better measure of the incorrectly classified cases than the Accuracy Metric.

$$Accuracy = \frac{TP + TN}{TP + FN + TN + FP}$$

Accuracy is the measure of all the correctly identified classes.

The TSC classification algorithm had an F1 score of 66.63%. From the diagram below, it can be seen that the time series classification algorithm had the highest accuracy and recall scores. The TSC model has an accuracy score of 91.2%, a recall score of 84.7% and a precision score of 97.3%. This means that the about 91% of the time, the model is able to predict the houses correctly. The MLP model had a very low recall score of about 0.2% which shows that it failed to correctly classify almost all the false negatives.

The TSC model performed better than the MLP model because the TSC model related the features to the time they occurred. This time feature helped the model in correctly predicting the houses.

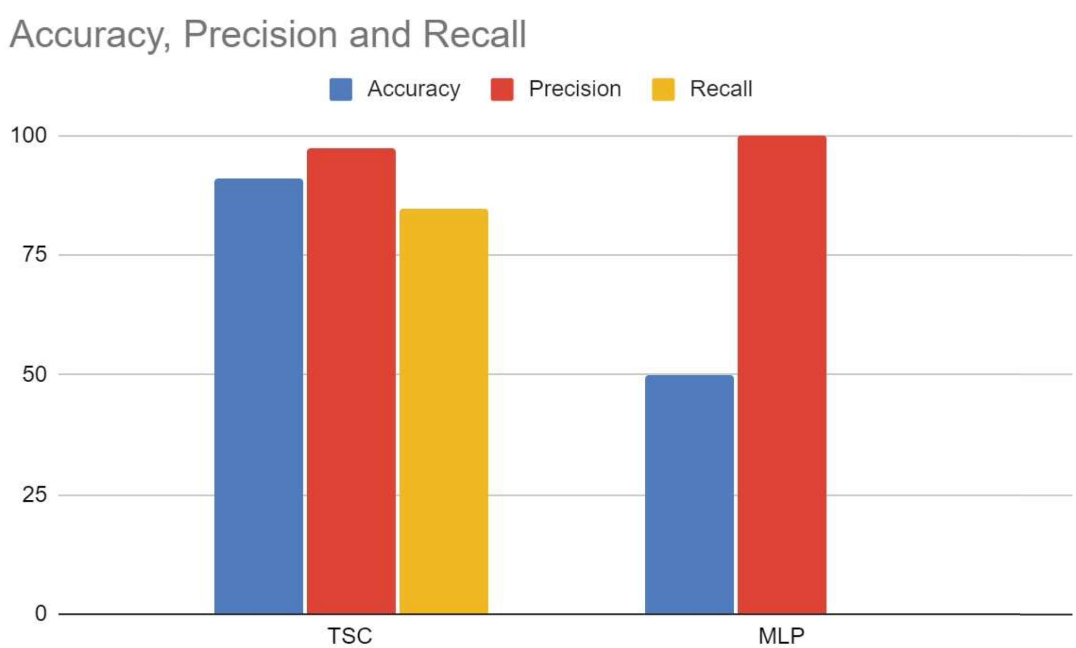

Figure 5.1 Output data in percentage

The table below summarizes the metrics used and the output values for both the TSC model and the MLP model.

Table 5.2 Summary of performance metric for both algorithms

Algorithms	Recall	Precision	Accuracy
TSC	84.7%	97.3%	91.2%
MLP	0.2%	100%	50%

From table 5.2, TSC has a better accuracy percentage than MLP. This analysis proves that the Time Series Classification Algorithm performs better at predicting the appliances and the tenants the belong to as compared to the Multilayer Perceptron.

Although, TSC is more efficient in distinguishing between similar appliances, it was extremely difficulty in predicting because of the inconsistency in the different timestamps at which data was collected and therefore in some cases made wrong predictions. Once, there is a consistent time interval in the collection of data, the possibility of more efficiently predicting the appropriate houses is high.

Chapter 6: Conclusion

6.1 Conclusion

To investigate how to distinguish similar appliances used by different tenants and hence splitting bills to ensure fairness, an NILM energy monitoring system was implemented using Multilayer Perceptrons (MLPs). This capstone project compared this algorithm to the Time Series Classification algorithm to determine which algorithms performs better. In order to assess performance, both algorithms were trained with data from House 1 and House 2 in the Reference Energy Disaggregation Dataset (REDD). Out of the several appliances found in these houses, light, refrigerator and microwave were chosen. From measuring, the performance using the Confusion matrix, it is concluded that the proposed Time Series Classification Algorithm (TSC) performs better than the Multilayer Perceptron (MLP) hence being more appropraite for distinguishing between similar appliances used by tenants.

6.2 Challenges Faced

In as much as this project was a success, there were a few challenges. The late arrival of the some components made it impossible to collect enough data for better analysis especially regarding the Ghanaian context. An implementation for an ideal compound house was not possible.

The available microcontroller, ATMEGA 2560 could not support the algorithms. Since the timing patterns of the different appliances were analyzed, it was necessary to have a large amount of data for a long period of time. It is computationally expensive to host such an algorithm on the microprocessor as it has low processing power.

The REDD dataset did not have a consistent time series. The power consumption data collected was not done at equal time intervals and so it was difficult to achieve high precision and accuracy when training and testing the data with both the Time Series Classification algorithm (TSC) and Multilayer Perceptron (MLP).

6.3 Future Works and Recommendations

In future works, a local dataset for the power consumption of compound houses in Ghana would be built to enhance further research in this field of energy disaggregation. The entire performance of the system would be evaluated, therefore the proposed system would be fully implemented in hardware as well as well as data collection over a long period of time. The use of a power adapter in the system would increase the possibility of identifying several features that can be used in load identification.

To be able to run the machine learning algorithms over Arduino, a recommendation would be to use an Arduino nano 33 BLE which is able to do machine Learning using Tiny ML. It runs at 64MHz with 1MB Flash memory and 256KB of RAM. It is well adapted for machine learning on constrained devices.

Data collection would be collected at consistent time intervals. This would make prediction and accuracy much better. Another recommendation with regards to achieving an improved accuracy would be to avoid splitting the data into training and testing in csv and just have one full dataset. Also adding more appliances to the system improves the relation between the data. Using a few number of appliances makes the data closely related. This makes it difficult for the algorithm make the separation.

References

[1] Mehta, *Principles of Power Systems*, vol. 53, no. 9. 2013.

[2] M. A. Toman and B. Jemelkova, “Energy and Economic Development: An Assessment of the State of Knowledge,” *Energy J.*, vol. 24, no. 4, pp. 93–112, 2003, doi: 10.5547/ISSN0195-6574-EJ-Vol24-No4-5.

[3] S. Kimathi Kuenyehia, S. Odame, and P. Arde-Acquah, “Electricity regulation in Ghana,” 2019, [Online]. Available: https://uk.practicallaw.thomsonreuters.com/w-019-3060?transitionType=Default&contextData=(sc.Default)&firstPage=true&bhcp=1.

[4] L. Asante, A. Sasu, J. Ayitey, and E. Gavu, “the Changing Face of Compound Houses in Ghana and Its Effect on Rental Value: a Case Study of Selected Neighborhoods in Kumasi, Ghana,” no. 2007, pp. 128–148, 2019, doi: 10.15396/afres2015_138.

[5] S. O. Afram and S. E. Owusu, “Design Innovations Towards enhancing the quality of living in multi-storey compund housing for low-income households in Kumasi, Ghana,” Journal of Science and Technology, Kumasi, 2006.

[6] R. A. Saaka, “A SMART ENERGY MONITOR FOR COMPOUND HOUSES THAT USE ONE SHARED METER,” *Ashesi Univ. Coll.*, vol. 1, no. 1, p. 47, 2019, [Online]. Available: http://www.ghbook.ir/index.php?name=فرهنگ و رسانه های نوین&option=com_dbook&task=readonline&book_id=13650&page=73&chkhashk=ED9C9491B4&Itemid=218&lang=fa&tmpl=component%0Ahttp://www.albayan.ae%0Ahttps://scholar.google.co.id/scholar?hl=en&q=APLIKASI+PENGENA.

[7] F. Babereyir, “THE COST- EFFECTIVENESS OF NED PROVIDING METERS TO EACH TENANT IN COMPOUND HOUSES IN NORTHERN AREA,” Kwame Nkrumah University of Science and Technology, 2011.

[8] L. Massidda, M. Marrocu, and S. Manca, “Non-intrusive load disaggregation by convolutional neural network and multilabel classification,” *Appl. Sci.*, vol. 10, no. 4, 2020, doi: 10.3390/app10041454.

[9] T. Bernard Dr.-Ing., “Non-Intrusive Load Monitoring (NILM): combining multiple distinct electrical features and unsupervised machine learning techniques,” 2018, [Online]. Available: https://duepublico2.uni-due.de/rsc/thumbnail/duepublico_mods_00046575.png.

[10] C. Peng, G. Lin, S. Zhai, Y. Ding, and G. He, “Non-Intrusive Load Monitoring via Deep Learning Based User Model and Appliance Group Model,” *Energies*, vol. 13, no. 21, p. 5629, 2020, doi: 10.3390/en13215629.

[11] R. Bonfigli, S. Squartini, M. Fagiani, and F. Piazza, “Unsupervised algorithms for non-intrusive load monitoring: An up-to-date overview,” *2015 IEEE 15th Int. Conf. Environ. Electr. Eng. EEEIC 2015 - Conf. Proc.*, pp. 1175–1180, 2015, doi: 10.1109/EEEIC.2015.7165334.

[12] A. Ruano, A. Hernandez, J. Ureña, M. Ruano, and J. Garcia, “NILM techniques for intelligent home energy management and ambient assisted living: A review,” *Energies*, vol. 12, no. 11, pp. 1–29, 2019, doi: 10.3390/en12112203.

[13] G. W. Hart, “Nonintrusive Appliance Load Monitoring,” *Proc. IEEE*, vol. 80, no. 12, pp. 1870–1891, 1992, doi: 10.1109/5.192069.

[14] Sense, “I Have Two of the Same Devices . Will Sense Know the Difference ?,” 2018. .

[15] S. Gupta, M. S. Reynolds, and S. N. Patel, “ElectriSense: Single-point sensing using EMI for electrical event detection and classification in the home,” *UbiComp’10 - Proc. 2010 ACM Conf. Ubiquitous Comput.*, pp. 139–148, 2010, doi: 10.1145/1864349.1864375.

[16] F. De la Prieta *et al.*, “Trends in Cyber-Physical Multi-Agent Systems. The PAAMS Collection - 15th International Conference, PAAMS 2017,” vol. 619, no. September 2017, 2017, doi: 10.1007/978-3-319-61578-3.

[17] A. Conesa, R. Bargallo, G. Velasco, and M. Roman, “Burden resistor selection in current transformers for low power applications,” *Semin. Anu. Automática, Electrónica Ind. e Instrumentación*, no. 1, pp. 5002–5005, 2014, [Online]. Available: http://www.saaei.org/edicion14/.
[18] M. Stusek, J. Pokorny, P. Masek, J. Hajny, and J. Hosek, “A non-invasive electricity measurement within the smart grid landscape: Arduino-based visualization platform for IoT,” *Int. Congr. Ultra Mod. Telecommun. Control Syst. Work.*, vol. 2017-Novem, pp. 423–429, 2017, doi: 10.1109/ICUMT.2017.8255171.
[19] T. Gupta, “Performance Measures for Classification,” 2020. https://towardsdatascience.com/performance-measures-for-classification-models-a486c8976bf1.

Appendix

Appendix A: Questionnaire

1. Who is your electricity provider?

2. Who makes the payment of electricity bills?

3. How many shared meters are on the compound?

4. How many people share one meter?

5. What are the issues that arise when it is time to share the bill?

6. How are the payments done since you use a shared meter?

7. How often do you pay your bills?

8. What is the monthly bill you receive in a year as a result of consumption?

9. What would you suggest should the solution to this billing issue?

Appendix B: Responses from tenants

Response 1 date: 17th January, 2021

Here is a response I received from a tenant in a compound house at Dome Pillar 2 which is the Greater Accra Region.

1. Who is your electricity provider?

Answer: The Electricity Company of Ghana

2. Who makes the payment of electricity bills?

Answer: Myself and the other tenant. It's a pre-paid meter and we make the purchase

3. How many shared meters are on the compound?

Answer: On this compound it is just one. The others use individual meters

4. How many people share one meter?

Answer: Two of us share this meter

5. What are the issues that arise when it is time to share the bill?

Answer: There are always issues of how much each person should pay because we consume power differently.

6. How are the payments done since you use a shared meter?

Answer: We have decided to pay 20 cedis each every week. So what happens is that when one person buys 20 cedis and it gets finished then another person buys 20 cedis.

7. How often do you pay your bills?

Answer: We buy pre-paid every week

8. What is the monthly bill you receive in a year as a result of consumption

Answer: To estimate the total amount, we would say 160 – 200 cedis.

9. What would you suggest should the solution to this billing issue?

Answer: A system that would ensure that no one is cheated, where billing would be done based on who consumes more.

One tenant mentioned that he uses only lights and an oven whereas the other tenant used an oven, lights and a fridge as well and may even be using other appliances on known to him because within a short time the power bought is consumed. He is thinking about sleeping in darkness so he does not pay at all because he feels cheated.

Response 2 date: 17th February, 2021

Here are responses I received from about 10 tenants in 7 different compound houses (of which 4 compounded houses had a minimum of 5 tenants) at Madina which is the Greater Accra Region.

1. Who is your electricity provider?

Answer: The Electricity Company of Ghana

Since there are several responses, I would give a summary.

Out of these 7 compound houses, 4 used post-paid meters and 3 used pre-paid meters. Most complaints about electricity billing had to do with the houses that used pre-paid meters. One major complaint was that there are always issues on who should pay or contribute when the credit is exhausted and as a result some of these tenants have purchased their own meters to avoid such issues. Another concern was an accusation of some tenants consuming more power than other because they assume some household members use appliances unknown to the other tenants and this results in the pre-paid credit being exhausted within a short period of time. Some tenants suggested the Electricity Company of Ghana provide a system where they know how much each tenant consumes.

Appendix C: Code for pre-processing (This involves the codes for; labelling of data, creating dataset for training and testing data)

```
# -----> DATA PROCESSING AND EXPLORATION


# @brief - This function attaches class labels values to dataset
# @param csv_file
# @param output_file
# @param new_file
# @param dataset
```

```
def append_output(csv_file, output, new_file):
    dataset = pd.read_csv(csv_file,parse_dates=['date']) # read old csv dataset
    dataset["Class"] = output  # attach output colum and output value
#     dataset.to_csv(new_file, index=False) # write results to a new file
    return dataset


# Append class labels for house 1 and 2 respectively
house1_dataset = append_output("house1_dataset.csv","1", "appended_house_1.csv")
house2_dataset = append_output("house2_dataset.csv","2", "appended_house_1.csv")
house1_dataset.head()


# Select Revelant Columns for each house
house1_dataset =  pd.concat([house1_dataset['timestamp'],house1_dataset['refrigerator'],
house1_dataset['light'], house1_dataset['microwave'],house1_dataset['Class']] , axis=1)
house2_dataset =  pd.concat([house2_dataset['timestamp'],house2_dataset['refrigerator'],
house2_dataset['light'], house2_dataset['microwave'],house2_dataset['Class']] ,


# ------> GROUP DATA FILES FOR TRAINING AND TESTING

# Select 500000 for training dataset and leave the rest for test
# Group for Training
house1_train = pd.concat([house1_dataset[0:125000],house1_dataset[-125000:]], axis=0) #
Slicing 250000 for training
house1_test = house1_dataset[125001:175001]

# Group for Testing
house2_train = pd.concat([house2_dataset[0:125000],house2_dataset[-125000:]], axis=0) #
Slicing 250000 for training
house2_test = house2_dataset[125001:175001]

house1_train.head()




# -----> EXPORT TO CSV
# Train Data
full_train_dataset = pd.concat([ house1_train, house2_train ], axis=0)
full_train_dataset.to_csv('full_train_dataset.csv' , index=False)

# Test Data
full_test_dataset = pd.concat([ house1_test, house2_test], axis=0)
full_test_dataset.to_csv('full_test_dataset.csv' , index=False)
```

In [75]:

```
# -----> LOAD FULL DATASET
```

```
file = tf.keras.utils
raw_df = pd.read_csv('full_train_dataset.csv')
raw_df.head()

# Pandas is a Python library with many helpful utilities for loading and working with
structured data
# and can be used to download CSVs into a dataframe.
```

Appendix D: Code for cleaning, splitting and normalize data

```
# ------> ClEAN , SPLIT, NORMALIZE

cleaned_df = raw_df.copy()
cleaned_df

# Use a utility from sklearn to split and shuffle our dataset.
train_df, test_df = train_test_split(cleaned_df, test_size=0.2)
train_df, val_df = train_test_split(train_df, test_size=0.2)

# Form np arrays of labels and features.
train_labels = np.array(train_df.pop('Class'))
bool_train_labels = train_labels != 0
val_labels = np.array(val_df.pop('Class'))
test_labels = np.array(test_df.pop('Class'))

# Features arrays
train_features = np.array(train_df)
val_features = np.array(val_df)
test_features = np.array(test_df)
                                                                    In [80]:
# Normalize the input features using the sklearn StandardScaler.
# This will set the mean to 0 and standard deviation to 1.

scaler = StandardScaler()
train_features = scaler.fit_transform(train_features)

val_features = scaler.transform(val_features)
test_features = scaler.transform(test_features)

train_features = np.clip(train_features, -5, 5)
val_features = np.clip(val_features, -5, 5)
test_features = np.clip(test_features, -5, 5)

print('Training labels shape:', train_labels.shape)
print('Validation labels shape:', val_labels.shape)
print('Test labels shape:', test_labels.shape)

print('Training features shape:', train_features.shape)
print('Validation features shape:', val_features.shape)
```

```
print('Test features shape:', test_features.shape)
```

Appendix E: Run test and label predictions codes

```
# -------> RUN TESTS AND LABEL PREDICTIONS

# model.predict(train_features[:10])
# model.predict_classes(train_features[:10])



# model.predict_classes(test_features[:10])
# test_features[0:10]
# test_labels[0:10]

predictions = model.predict(test_features[:10])
score = tf.nn.softmax(predictions[9])

print(test_features[:10])
print(test_labels[:10])
print("  ")
print(" Prediction ")
print(predictions)

print("  ")
print(" Score ")
print(score)
print(test_labels[np.argmax(score)])

print("  ")
print(" Predict Classes ")
print(model.predict_classes(test_features[:10]))
label = predictions.argmax(axis=-1)
print(label)

# print(test_labels[score])

# # print(test_features.class_names)

print(
    "This consumtion data most likely belongs to House {} with a {:.2f} percent confidence."
    .format(test_labels[np.argmax(score)], 100 * np.max(score))
)
```

Appendix F: Code for Evaluate Model with Metrics

```
-------> EVALUATE MODEL WITH METRICS


train_predictions_baseline = model.predict(train_features, batch_size=BATCH_SIZE)
```

```
test_predictions_baseline = model.predict(test_features, batch_size=BATCH_SIZE)


def plot_cm(labels, predictions, p=0.5):
  cm = confusion_matrix(labels, predictions > p)
  plt.figure(figsize=(5,5))
  sns.heatmap(cm, annot=True, fmt="d")
  plt.title('Confusion matrix @{:.2f}'.format(p))
  plt.ylabel('Actual label')
  plt.xlabel('Predicted label')

  print('Legitimate Transactions Detected (True Negatives): ', cm[0][0])
  print('Legitimate Transactions Incorrectly Detected (False Positives): ', cm[0][1])
  print('Fraudulent Transactions Missed (False Negatives): ', cm[1][0])
  print('Fraudulent Transactions Detected (True Positives): ', cm[1][1])
  print('Total Fraudulent Transactions: ', np.sum(cm[1]))


baseline_results = model.evaluate(test_features, test_labels,
                                  batch_size=BATCH_SIZE, verbose=0)
for name, value in zip(model.metrics_names, baseline_results):
  print(name, ': ', value)
print()

plot_cm(test_labels, test_predictions_baseline)
```

Appendix G: Codes to Graph timestamps against the appliances

```
# # # # # GRAPH PLOT FOR TIMESTAMP AND REFRIGERATOR
import matplotlib.pyplot as plt

appended_house1 = appended_house1[10000:200000]
appended_house2 = appended_house2[10000:200000]
appended_house3 = appended_house3[10000:200000]

df1 = pd.DataFrame(appended_house1, columns=['timestapm','refrigerator'])
df2 = pd.DataFrame(appended_house2, columns=['timestapm','refrigerator'])
df3 = pd.DataFrame(appended_house3, columns=['timestapm','refrigerator'])

df1 = df1.cumsum()
df2 = df2.cumsum()
df3 = df3.cumsum()

plt.plot( 'timestapm', 'refrigerator', data=df1, marker='', color='blue',label='house1')
plt.plot( 'timestapm', 'refrigerator', data=df2,marker='', color='red',label='house2')
# plt.plot( 'timestapm', 'refrigerator', data=df3, marker='', color='yellow',label='house3')

plt.xlabel('timestamp')
plt.ylabel('refrigerator')
```

```
plt.savefig('refrigerator.png')

plt.legend()
plt.show()


# In[20]:


# # # # # GRAPH PLOT FOR TIMESTAMP AND LIGHT
import matplotlib.pyplot as plt

appended_house1 = appended_house1[10000:200000]
appended_house2 = appended_house2[10000:200000]
appended_house3 = appended_house3[10000:200000]

df1 = pd.DataFrame(appended_house1, columns=['timestapm','light'])
df2 = pd.DataFrame(appended_house2, columns=['timestapm','light'])
df3 = pd.DataFrame(appended_house3, columns=['timestapm','light'])

df1 = df1.cumsum()
df2 = df2.cumsum()
df3 = df3.cumsum()

plt.plot( 'timestapm', 'light', data=df1, marker='', color='blue',label='house1')
plt.plot( 'timestapm', 'light', data=df2,marker='', color='red',label='house2')
# plt.plot( 'timestapm', 'light', data=df3, marker='', color='yellow',label='house3')

plt.xlabel('timestamp')
plt.ylabel('light')

plt.legend()
plt.show()


# In[ ]:


# # # # # GRAPH PLOT FOR TIMESTAMP AND MICROWAVE

import matplotlib.pyplot as plt

appended_house1 = appended_house1[10000:200000]
appended_house2 = appended_house2[10000:200000]
appended_house3 = appended_house3[10000:200000]

df1 = pd.DataFrame(appended_house1, columns=['timestapm','microwave'])
df2 = pd.DataFrame(appended_house2, columns=['timestapm','microwave'])
df3 = pd.DataFrame(appended_house3, columns=['timestapm','microwave'])
```

```
df1 = df1.cumsum()
df2 = df2.cumsum()
df3 = df3.cumsum()

plt.plot( 'timestapm', 'microwave', data=df1, marker='', color='blue',label='house1')
plt.plot( 'timestapm', 'microwave', data=df2,marker='', color='red',label='house2')
# plt.plot( 'timestapm', 'microwave', data=df3, marker='', color='yellow',label='house3')

plt.xlabel('timestamp')
plt.ylabel('microwave')

plt.legend()
plt.show()
```