APPROVAL SHEET

Title of Thesis: **A Holistic Cyber-Physical System to Detect Anomalies in Smart Home Appliances**

Name of Candidate: Sarthak Pathak
Master of Science, 2017

Thesis and Abstract Approved:

Dr. Nirmalya Roy
Assistant Professor
Department of Information Systems

Date Approved: May 27, 2017

ABSTRACT

Title of Document: **A HOLISTIC CYBER-PHYSICAL SYSTEM TO DETECT ANOMALIES IN SMART HOME APPLIANCES**

Sarthak Pathak, Master of Science in Information Systems, 2017

Directed By: Assistant Professor, Nirmalya Roy, Department of Information Systems

This thesis focuses on detecting the anomalies of appliances using their acoustic signature and power consumption data. In this work, we advocate a holistic cyber-physical system (CPS) which helps recognize the sound and energy patterns generated by a variety of everyday appliances and provide better monitoring of them. We focus on the appliance health estimation by combining its acoustic signal and power consumption data, provide just-in-time feedback to the customer, and encourage them to take appropriate precaution to avoid any increase in electricity bill or any physical damage or loss. As every appliance creates a unique sound which is not audible to human ears, we captured those different frequencies sound which provide a lot of knowledge about the appliance's working condition. This system eliminates the potential risk of various types of sensor failure which are used in a smart home. Finally, we correlated and established the relationship between appliance's power consumption and acoustic signature. We extracted the features like fast Zero Crossing Rate (ZCR), Audio Energy, Energy Entropy, Spectral Centroid, Spectral Spread, Fourier transform (FFT), Mel-frequency Cepstral Coefficient (MFCC) from acoustic signal. We trained and tested the model with different classifiers such as support vector machine, ensemble, J48 decision tree, OneR and showed up to 90% accuracy.

A HOLISTIC CYBER-PHYSICAL SYSTEM TO DETECT ANOMALIES IN SMART HOME APPLIANCES

By

Sarthak Pathak

Thesis submitted to the Faculty of the Graduate School of the

University of Maryland, Baltimore County, in partial fulfillment

of the requirements for the degree of

Master of Science

2017

© Copyright by
Sarthak Pathak
2017

Acknowledgements

I would first like to thank my thesis advisor Dr. Nirmalya Roy of the Information Systems at University of Maryland, Baltimore County. The door to Dr. Roy office was always open whenever I ran into a trouble spot or had a question about my research or writing. He consistently allowed this project to be my own work, but steered me in the right direction whenever he thought I needed it. Dr. Roy has the right attitude and substance of a genius, he continually and convincingly conveyed a spirit of adventure regarding research and an excitement in regard to teaching. Without his guidance and persistent help this thesis would not have been possible.

I would like to thank my committee members, Dr. Dongsong, Zhang and Dr. George Karabatis for taking out time to be my committee member. I attended Dr. Zhang lectures and his teaching style and enthusiasm for the machine learning made a strong impression on me and I have carried and applied the knowledge that I gained from his teaching. Gratitude to Dr. Karabatis for helping me to form a committee.

I would like to thank all PhD students from Mobile Pervasive & Sensor Computing lab. I would like to thank Nilavra Pathak for suggesting a good project idea. I must thank Hafiz and Sajjad for constantly and passionately helping me on my several issues while working on this thesis. Working on Matlab could not be a fun without any help from Sreeni. He was the one with an honest feedback every time. I would like to thank Abu Zaheer who offered his encouragement through phone calls and emails every day. He has been kind and supportive to me over the last 9 months.

Most importantly, none of this could have happened without my friend’s support in collecting data from different homes.

Table of Contents

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

The Internet has made the world "level" by rising above space. We can now collaborate with individuals and get helpful data around the world in a small amount of a moment. The Internet has changed how we lead explore, considers, business, administrations, and amusement. Be that as it may, there is yet a genuine crevice between the digital world, where data is traded and changed, and the physical world in which we live. The developing cyber physical system frameworks should empower a cutting edge amazing vision for societal-level administrations that rise above space and time at scales never conceivable. Maintenance, operations, manufacturing, and design teams are under tremendous pressure to maximize appliance life and utilization without compromising safety and operational uptime. Most homes have number of appliances that are major energy users that typically consumes great power without being noticed. Certain appliances (such as laundry machine, drier, water heater, microwave, refrigerator) can lead to high consumption as well. Like older kitchen appliances cost between $90 and $350 a year to run and they can be shut down any time, so it is necessary to design, better, affordable and high-efficient monitoring system. 26% American homes have two (or more) fridges, 96% have a microwave, 90% have a stove, 82% have a clothes washer, 79% have a clothes dryer, 60% have a coffee maker and 59% have a dishwasher. Because of high use of appliances, accurate health assessment of appliance is one of the key element which enable predictive maintenance strategy to increase productivity, reduce maintenance costs and mitigate safety risk. For consumers, detecting anomaly in an appliance helps them in monitoring the health of an appliance. For example, a microwave which is been used for 5-6 years, we didn’t notice its power consumption increased by the time and one day some internal wiring fault causes to shut down whole house electricity. It also can damage other appliances or devices plugged in. To avoid these scenarios, we have designed system to monitor health of an appliances by detecting an anomaly in it. The main motivation for this technology is track and identify the working of an appliance in the most discrete way possible and preserving the privacy of the individual human. Beyond real-time data collection, we also aim to do frequency domain and time domain analysis on the sound signal to get an insight from the data and combining all features into proper feedback or recommendations. Peculiarity or anomaly recognition alludes to the issue of discovering examples in information that don't adjust to expected conduct. Patterns which are different and cannot be confirmed to be expected are referred to as outliers or anomalies. The biggest challenge for this system is to define and identify the correct pattern, which also work as a baseline. And, justify that anything observed other than base pattern is an anomaly. To know more about the sound signal, there are some sound signal analysis techniques performed like Fast Fourier Transform (FFT), Short Time Fourier Transform (STFT), Mel-frequency Cepstral Coefficient (MFCC) and several other features which defines sound signal more profoundly and many things information can be deduce from this little features. These features are very useful in area of Automatic Speech Recognition (ASR) and Crowd Sensing System (CSS).

1.1 Significance of the Problem

The U.S Energy Information Administration has been tracking appliances use by American because older appliance need a lot of energy to run. Biggest challenge is to know when they are going to fail? Residential buildings have multiple rooms with a variety of appliances and HVAC equipment. People spend thousands of dollars on maintenance, repair and servicing. Currently, appliance health checkup and maintenance performed only when complaint is made. It is not done on regular basis and sometimes it misses the crucial time for appliance health checkup. Continuous monitoring of appliances, early anomaly detection and prediction of failure before it happens can support maintenance strategies, prevent sudden downtime which result into increase in efficiency and decrease in cost. There are several techniques for appliance health checkup and maintenance that has been performed. Some of them are very expensive and time consuming too. For instance, the cost for replacing a damaged motor or part of an appliance can be around $1000 dollars or more. So, to deal with these significant problem, we have come up with some powerful system which uses microphone to sense the acoustic sound. Getting an acoustic sound, it is possible to infer, for example: coughing, emotions, stress, crowd density or number of people. Audio sensing is already in use in field of health and wellbeing where it monitors the user's health. Challenge for acoustic audio sensing is the diverse acoustic environment. Another challenge is to make an accurate predictions or inferences because each environment is so diverse and contains its own mixture of sounds that often confuses the audio classifiers. Quite a few "sensing urban soundscape" examples exist which measures the noise pollution in urban environments. They built a cyber-physical sensor network systems on citygram which measure, stream, analyze, visualize and infer urban soundscapes.

1.2 What is Anomaly Detection?

Anomaly is an abnormal pattern or behavior that we find. It can be find in any data set whether it is two dimensional or multi-dimensional data set. Suppose, in Figure 1, N1 and N2 are two observed normal regions. But then there is another region which is look little different, which can be considered as an anomaly or outlier. Now, this may be or may not be an anomaly, depending upon the context. For example, a microwave produces a beep before start and end of its operation. If we take acoustic data of microwave, then that beep sound can be an anomaly as it doesn't continuously beep. But here the context is different so we do not consider that beep sound as an anomaly. There are three types of anomaly: -

Point Anomaly: This is a simple type of anomaly where if an individual information occurrence can be considered as abnormal concerning whatever is left of information, then the occasion is named a point peculiarity or anomaly. For example, to detect credit card fraud, amount spent is considered as feature and through this fraud is detected.

Contextual Anomaly: An individual information is anomalous in a specific context, but it is not otherwise. It is also called conditional anomaly. For example, 32 degree Fahrenheit is considered as normal in winter, but this 32 degree Fahrenheit in Summer is considered as anomaly.

Collective Anomaly: Observed collection of data information is anomalous with respect to the entire data set. Occurrence of some instance in a group or in a collected fashion can be anomalous. For Example, differentiating two people voice is the perfect example for this. Sometime we are interested in one person voice so other person voice can be considered as an outlier, based on two people speaking frequency.

To detect these anomalies is called anomaly detection. Anomaly can be detected in supervised and unsupervised environment. Nature of input data is also very key aspect of anomaly detection. In real world application, use of anomaly detection has been increased. Domains such as intrusion detection, fraud detection, industrial damage detection, machine damage detection, anomaly detection in text area are the hot topics.

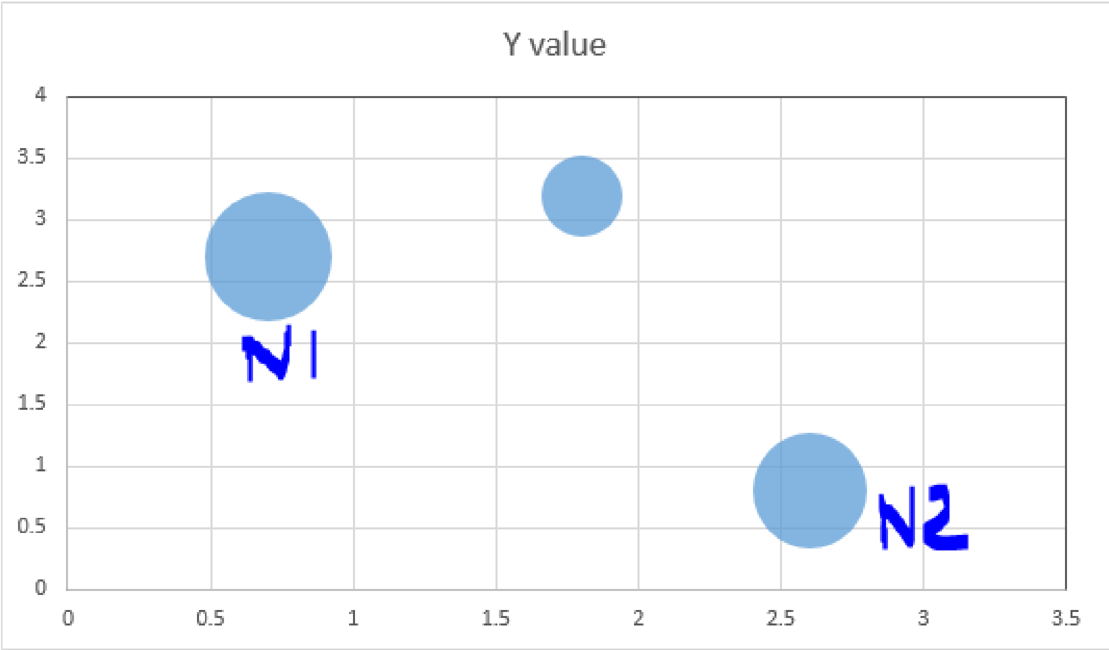


Figure 1 Anomaly Detection

1.3 Techniques used to detect anomaly

Once we have set of labeled data, we need that to train a model. We divide data set into training and testing set, as classification based anomaly detection works in two steps. First, training set(labeled) is used to train a classifier and then test data is used to test and model will classify whether is normal or anomaly. There can be one class classifier or multi-class classifier data set. Single class or one class classifier creates a boundary and any data instance found beyond that boundary is considered as anomaly. Classification techniques can be neural network based, Bayesian network based, support vector machines based, rule based. Other famous techniques used to detect anomaly are, nearest neighbor based anomaly detection, clustering, statistical based and spectral analysis based.

Chapter 2: Related Work

Our thesis was inspired by technologies which helped people in monitoring and maintaining the health of machines, appliances and helped people in detecting anomalies in real world. We will discuss all the related work done in the field of anomaly detection, sound signal processing, energy disaggregation, different classification algorithms used to detect anomalies.

Researchers like Bong Jun Ko, Jorge Ortiz, Theodoros, Dinesh Verma from IBM T.J Watson Research Center [1] architected a system which detect a fault in a machine room environment using acoustic sound of HVAC equipment, fans and generators, then processing them. Since fault only be found when it happens. So, this system continuously check and monitor the machine room environment and provide feedback to the person. They labeled the signal data manually as either "normal" or "abnormal" and run the ensemble machine learning algorithm to detect an anomaly. Work-order system is also integrated with this system which automatically gives the order to repair any equipment. The reason for using the ensemble learning is because it is observed that using multiple classifier works better than using single classifier. Trained classifier from two room acoustic sound is used for the third room. In the current experiment, they have extracted four features from a sound like Discrete Cosine Transform (DCT), multiple window Fast Fourier Transform (FFT), single window FFT and Mel-frequency Cepstral Coefficient (MFCC). In figure 2, it shown that labeled audio dataset is send to feature extractor where audio features are extracted and classifier engine is used to run classification algorithm on that feature file. Based on policy, the feedback is generated and root cause analysis has been done to it.

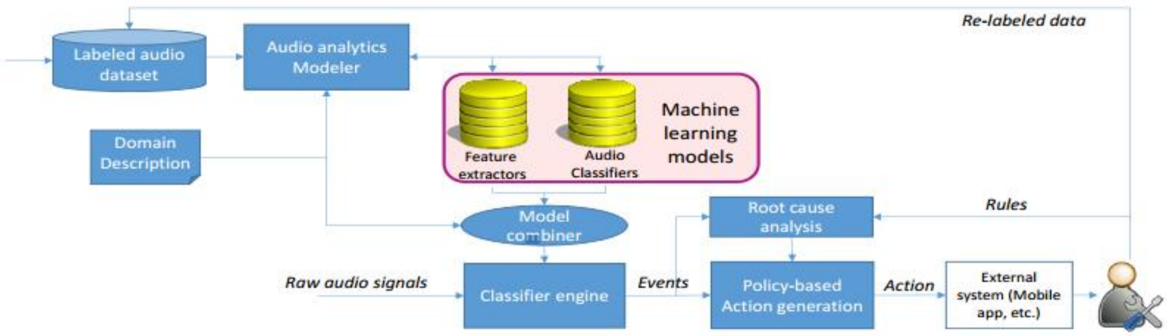


Figure 2 IBM anomaly detection in machine room through acoustic sound

Another similar approach taken in field of audio sensing by Nicholas D. Lane, Petko Georgiev and Lorena Qendro [2] from Bell Labs, University of Cambridge where they have developed audio sensing system using smart phone in unconstrained acoustic environment using deep learning. This is very unconventional approach using smart phone and achieve very high accuracy and greater robustness to distinguish ambient scene, recognize emotions and identify the speaker. Using multi-layer deep learning algorithm proves to be effective. They also prove that use of smart phone is economical and it also saves lot of battery. Even though training require lot of dataset and significant computation power.

EmotionSense: A mobile phone based adaptive platform for experimental social psychology research. They developed a system which has the ability of sensing individual emotions as well as activities, verbal and interaction between them. Emotions are detected and monitored using people's voice data which is continuously recorded by mobile phone. They have used Gaussian Mixture method for emotion detection and speaker identification. Classifiers are running locally on off the shelf mobile phones.

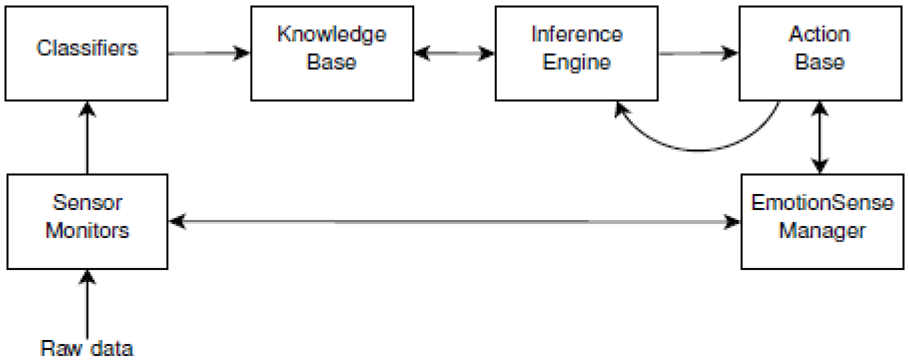


Figure 3 Audio sensing system by Bell Labs

Majority of work has been done in the field of sensing the urban or surrounding sound through microphone or smartphone and developing such kind of system which record, analyze and compute into it. The design of urban sound monitoring devices by Charlei Mydlarz, Samuel Nacach, Tae Honh Park and Agnieszka [3] designed an acoustic sensing system device where they incorporated android device with wi-fi, MEMS microphone and a USB audio device. As noise pollution is increasing threat, so to measure the noise pollution in the city, they have created this device which they named Tronsmart MK908ii, figure 4.

Figure 4 Tronsmart

Aims to continuously monitor and ultimately understand these urban sound environments. Urban sound monitoring helps in real estate values (city noise). They designed acoustic sensing devices with wireless communication called Remote Sensing Device(RSD). RSD = mini android based PC + MEMS microphone. Compared their device with commercial RSD (Libelium Waspmote) which cost around $ 700. They build a device which will cost around $ 70, they named it Tronsmart MK908ii Android based mini PC.

Same team also worked in the implementation of MEMS microphone for urban sound sensing [3] where they compare the frequency response and dynamics range of each system available in market, figure 5.

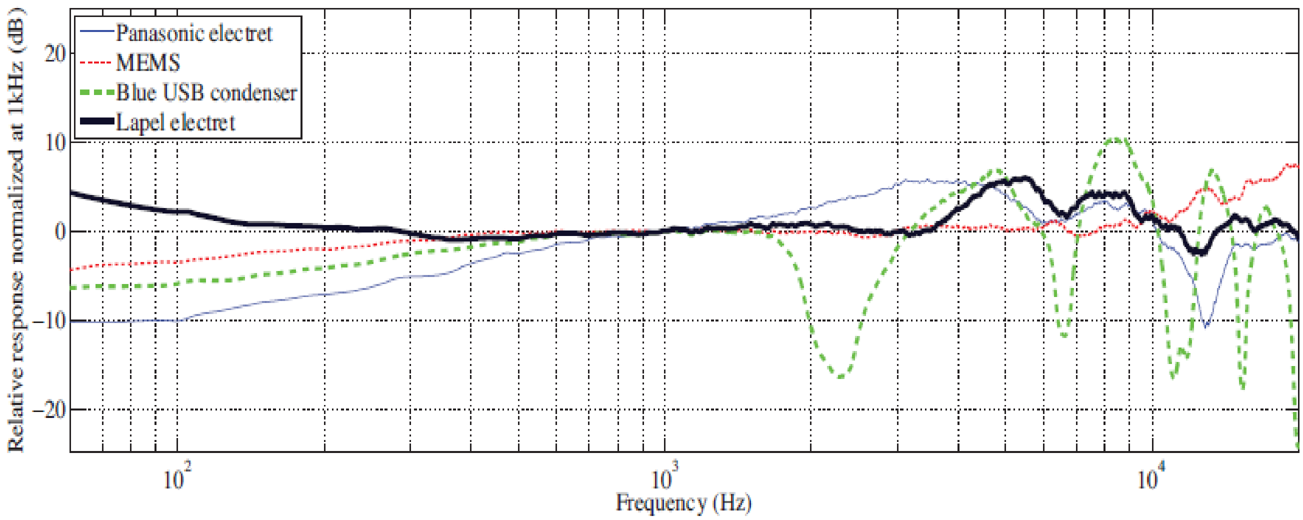
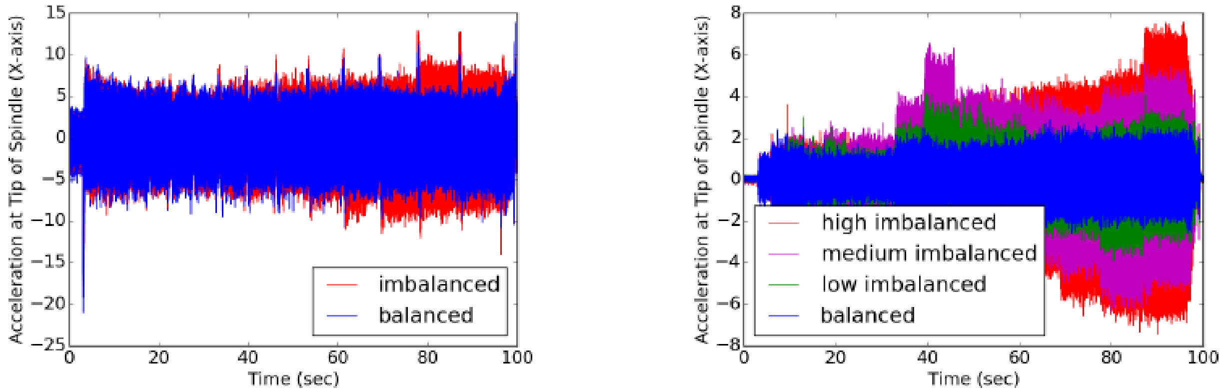


Figure 5 Comparison of different audio sensing devices

If we talk about monitoring devices or appliances, there are some work has been done where use of sensors make things easier and better. Linxia Liao, Tomonori Honda and Radhu Pavel [4] did some work in area of estimating device health where they combine the contextual information with help of sensors. They have also used restricted Boltzmann machine. There experiment shows that anomaly or imbalance can be accurately identify when we combine device operating environment context with sensor data. They analyzed sensor data and extracted features such as average, standard deviation, maximum amplitude of FFT and frequency for maximum amplitude of FFT. Figure 6 shows different timeline data. Used Restricted Boltzmann machine as a feature generation model and coupled with Random Forest algorithm for generating predictions.



(a) Spindle x-axis acceleration: 2009 data.

(b) Spindle a-axis acceleration: 2010 data.

Figure 6 Anomalous data comparison

Significant number of anomaly detection and diagnosis methods have been proposed for machine or device health monitoring. A survey on anomaly detection [5] by Varun Chandola, Arindam Banerjee and Vipin Kumar from University of Minnesota in year 2007 explains various techniques like classification, clustering, statistical, spectral analysis are grouped into different categories to differentiate between normal and anomalous behavior. The nearness of information crosswise over various appropriated areas has inspired the requirement for circulated peculiarity identification methods [Zimmermann and Mohay 2006]. While such procedures handle data accessible at different locales, they frequently need to all the while secure the data exhibit in each site, in this way requiring security saving abnormality identification strategies [Vaidya and Clifton 2004]. Another up and coming range where inconsistency recognition is discovering increasingly pertinence is in complex frameworks. A case of such a framework would be an air ship framework with various parts. Abnormality recognition in such frameworks includes displaying the cooperation among different parts [Bronstein et al. 2001].

MD Abdullah Al Hafiz et al. [6] devised and system a demo for a microphone sensor based system for green building applications where they have designed multi-modal energy disaggregation algorithms. With help of this they identified appliance state and efficiently controlled the HVAC system. They have used Raspberry Pi with super high gain microphone, figure 7, which they have deployed indoor and with use of Sound eXchange (SoX) script they have split the noise and human voice. Feature extracted is Fast Fourier transform (FFT). With help of this acoustic sensing device they efficiently identified state, abnormal power consumption, event recognition and detected occupants in building. Another aspect of this experiment is to measure the energy consumption by CPU based FFT computation vs GPU based energy consumption.

Figure 7 Raspi Audio sensing device

Chapter 3: Methodology

This chapter describes the method and architecture of the model that we have designed to collect acoustic sound and energy data from two different sensors and with help of that data how we extracted meaningful information about an appliance and run a classification algorithm on top of it. This section starts with explaining the overall high level architecture of the system followed by the data collection, feature extraction, spectral analysis and classification. We also investigated the ways to make a hybrid model by combining energy and acoustic data and checked how accurate is the model to classify normal and abnormal condition of appliances.

3.1 Architecture of the System

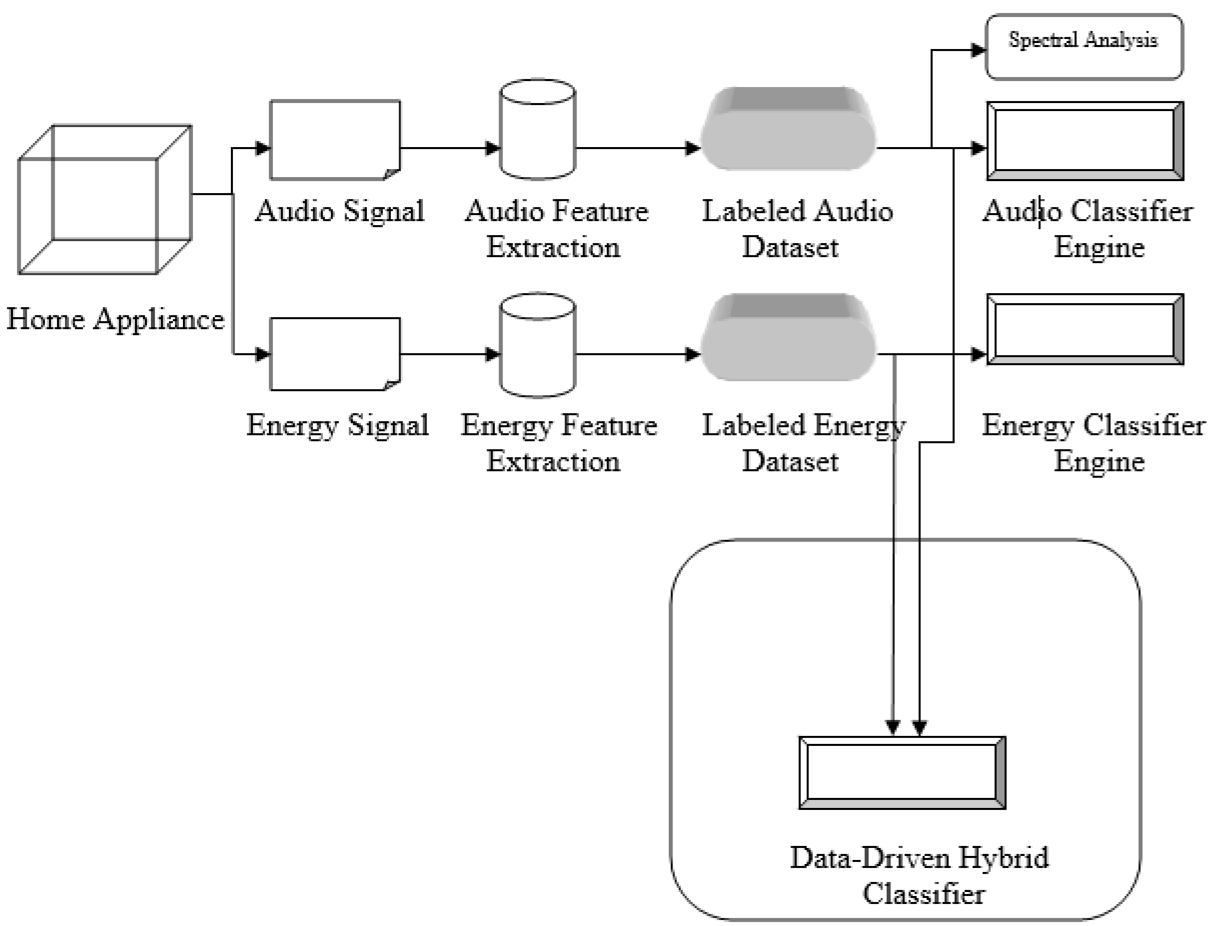


Figure 8 System Architecture

The system architecture shown in the above figure 8, gives the complete picture of the overall methodology. There are two types of data is collected from two different sensors. One is acoustic audio signal, which is collected from a microphone enabled mobile phone with a high gain. Another data is energy data which is collected by a smart plug. For our thesis, we have collected used an Enmetric smart plug [7], which collects the energy raw signal and sends it to the Enmetric cloud, Figure 9.

Figure 9 Enmetric Smart Plug

After collecting two different types of raw data from two different types of sensor, it is very important to extract some meaningful information from both types data. There are two feature separate extractors for both audio and energy. Audio feature extractor will extract 21 different features from audio. Some of the audio features are MFCC, Energy, Zero Crossing Rate, Spectral spread, Spectral Centroid etc. As every feature extracts, its own meaningful information such as spectral centroid gives the center of gravity of the spectrum. Entropy of Energy feature will measure some abrupt changes in audio. Energy feature extraction will extract some features like Minimum Power, Maximum Power, Power Frequency etc.

All features from both acoustic and energy will be labeled as normal or abnormal manually. Each labeled file will be send to classifier engine in which classification algorithm will be applied such as support vector machine, j48, IBK, OneR, ensemble. The reason to use ensemble because it is claimed that use of multiple classifier gives better result than use of single classifier. That model can also be used to classify unknown appliance abnormal or normal nature. Classifier engine is different for different inputs. Like for audio signal, classifier engine is different and it uses support vector machine but in case of energy data the classifier engine is different which uses random forest to classify normal and abnormal class. In addition to this spectral analysis is also perform on audio signals. Audio signal in time domain analysis and frequency domain analysis can be seen with different color codes explaining different frequencies with different energies at a given time, it is explained in details below. Another major part of this architecture is the model combiner or data-driven hybrid classifier. We would like to investigate on accuracy of system when acoustic data and energy data combines. We tried to merge energy and acoustic features into one. Both, acoustic and energy are having different resolution data. We merge them with equal resolution and created a single file having both features (acoustic and energy) into one. We applied different classification algorithm on that dataset. Bagging and boosting also applied, as the dataset is skewed and there might be loss of some information due to merging different resolution data. Finally, accuracy has been recorded from all classifier engines.

3.2 Data Collection

To perform all experiments, we have collected our own data from appliances. Primarily, we have collected all the data (acoustic and energy) from microwave. To collect data of different time interval, we have used microwave of 1-year-old, 2-year-old, 5-year-old, 25-year-old and newly purchased which is manufactured in year 2017. Notwithstanding this we likewise purchased flawed microwave 25-year-old from Goodwill [8] store. We considered this old microwave as flawed one because of certain criteria like energy consumption, internal physical damage, time taken to heat food or heat dissipation. Our aim to have these distinctive microwaves so that we can have an expansive range of information with which we can get more data and with help of this information we can better train our classifier model. There were many challenges to gather these broad spectrums of microwaves. So, we went to different homes and requested permission to collect theirs microwave acoustic and energy data. Through this approach we had 1-year-old, 2-year-old and 5-year-old microwave data. To get new microwave data, we purchased it from Walmart [9]. To train our model to classify into normal and abnormal appliance, it was important for us to have some abnormal instances. Fortunately, we found one flawed microwave from Goodwill store which is working but the motor which rotates the plate on which we keep food is not rotating. Moreover, energy consumption is higher in compared to other microwaves. Due to which we labeled it into an abnormal microwave.

3.2.1 Experimental Setup

The audio samples are recorded in clean environments without background noise. We have used Motorola droid turbo XT1254, Figure 10, microphone to record the audio samples. We recorded audio samples by placing microphone on two positions, one is on side and other one is on back of the microwave. Later, our research help us to identify proper microphone placement to capture an acoustic sound of appliance. To collect energy data, we have used Enmetric smart plug. The reason to use Enmetric is that it is natural and sparing measures agreeable equipment, gathers and conveys information to cloud data administrations. It is easy to deploy and scale to fit your needs. Enmetric monitors energy usage and it can extract features like Maximum Power, Minimum Power, Average Power, Power frequency. We collected both acoustic and energy data at same time to make some correlation between them. Energy data that we collected has resolution of 1 minute. We have created several test cases while collecting data. We collected 3 minutes of data every time. First, we collected data of empty microwave. Second, a microwave filled with food and a glass of water. We collected data from all the different microwaves but we make sure that we have all microwave of same power rating. As different power-rated microwave can give different results, for example 700 Watt microwave takes 1 minute to boil a water but 1200 Watt will take 40 seconds or less to boil a water. Here in all our experiments we focused on 700 Watt microwaves. All acoustic sounds are stored on cloud but it was processed on a system. But energy data was stored and processed on cloud. We can any time download and use energy data. Enmetric extracts all the features from energy data and provide a .CSV file. All acoustic data are WAV file, as it gives best result than any other format like MP3, which is compressed form.

Figure 10 Audio Sensor, Motorola Droid Turbo XT1245

Figure 11 Experimental Setup, 25-year-old microwave and Enmetric device

3.3 Feature Extraction

Feature extraction is the most important part in our system. After collecting data, it is the first step of processing. In this section, we will explain about sound signal processing for machine learning which plays the most important role in any audio detection or speech recognition system. We wrote python script to extract audio features. We will further explain several features that we have extracted from audio and energy.

3.3.1 Audio Features:

In this sort of experiment extricating information from sound is extremely urgent and to pick which feature gives better outcome is somewhat the primary concern to do. Several audio features both from time and frequency domain are extracted. The complete detailed list of features is presented in table [1]. The time domain features are directly extracted from the raw file. As the raw audio file is represented in a time domain Figure 12. This figure is direct represented in time domain from raw audio file. We plot a graph of time against amplitude. The graph represents at a given time what is the power or amplitude of the sound signal. But this representation cannot give lot of information.

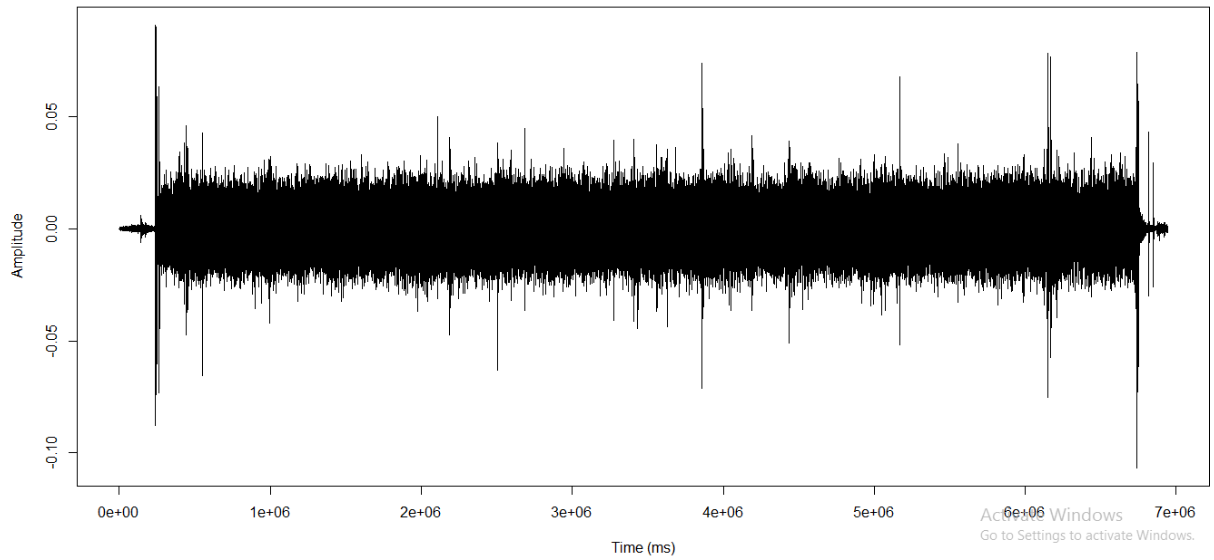


Figure 12 Time domain representation of sound signal

To extract features we split whole audio is short time frames also called windows. Widely accepted frame size is 20-100ms. Which is the resolution of the audio file. After splitting into frames, we conducted further process by overlapping frames to get an information of a millisecond of a data. In overlapping we kept frame step size shorter than frame length. All our features are calculated from these frames. After calculating all the features, we stored them into a .CSV file. Both time domain and frequency domain feature vectors are stored in a single file. This is how we created feature file for all the audio signals. First we extracted three features form raw sound file on time domain. Features like Zero Crossing Rate (ZCR), Energy and Entropy of Energy are the only features from time domain. Zero Crossing Rate (ZCR) is the rate of sign changes of the audio signal from positive to negative or vice-versa during the duration of the specific frame. Energy is the sum of the squares of the signal values. Entropy of Energy measure the abrupt changes in the signal of each frames. We also plot some graph showing the ZCR and Energy on time domain. Figure 13, shows a new microwave of 700 watts filled with a food. Figure 14, shows a ZCR and Energy graph of a 5-year-old microwave with food.

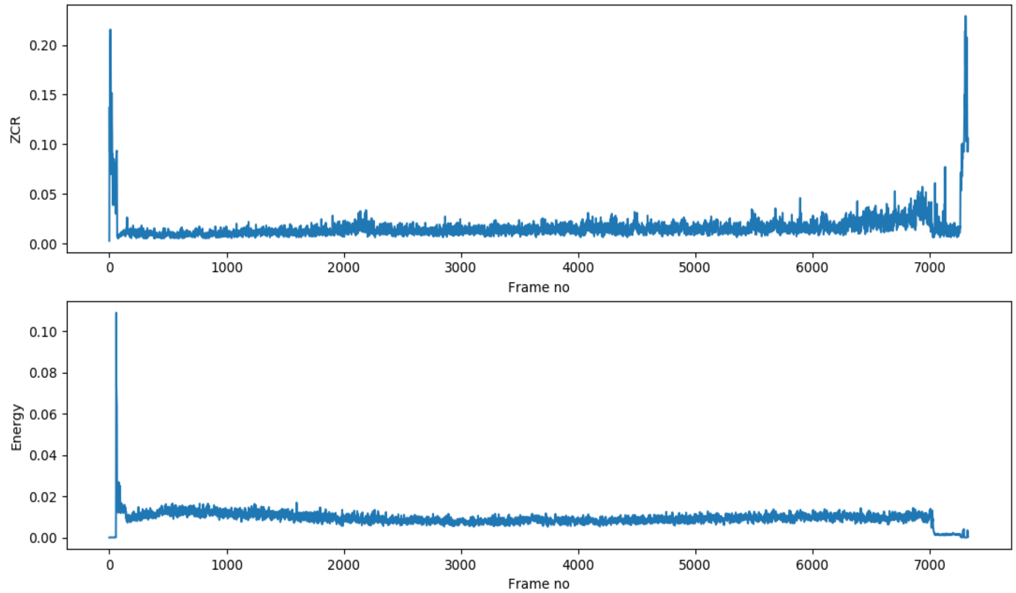


Figure 13 New microwave time domain features

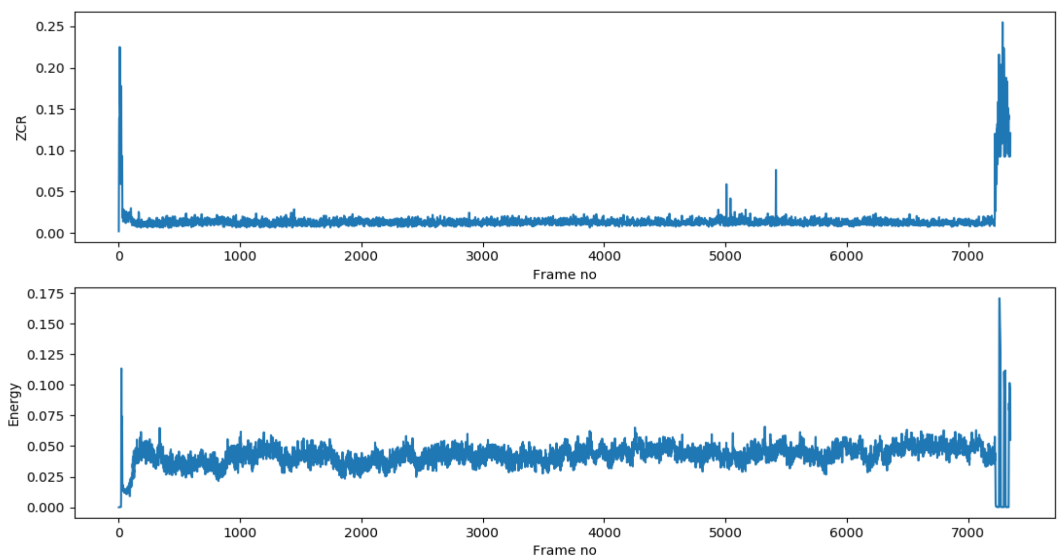


Figure 14 5-year-old microwave time domain features

Other important feature that we have extracted from frequency domain of sound signal is Mel-Frequency Cepstral Coefficient (MFCC). The fundamental indicate comprehend about discourse is that the sounds created by a human are separated by the state of the vocal tract including tongue, teeth and so on. This shape figures out what sound turns out. On the off chance that we can decide the shape precisely, this ought to give us an exact portrayal of the phoneme being created. The state of the vocal tract shows itself in the envelope of the brief span control range, and the employment of MFCCs is to precisely speak to this envelope. Mel Frequency Cepstral Coefficients (MFCCs) are an element broadly utilized as a part of programmed discourse and speaker acknowledgment. They were presented by Davis and Murmelstein in the 1980's, and have been best in class from that point onward. Preceding the discovery of MFCCs, Linear Prediction Coefficients (LPCs) and Linear Prediction Cepstral Coefficients (LPCCs) were the principle include sort for Automatic Speech Recognition (ASR), particularly with HMM classifiers. MFCC is implemented in 6 steps, Figure 15. At the end, we get a matrix containing the 13 MFCC vectors for each processed window. For this experiment, we have used window size configuration of 25ms length and 10ms frame step size. So, we got matrix size of 13x(100*s) for an s seconds. MFCC's are good for the lower frequency audio data and in our case, most of the data that we have is in the lower frequency. Speech related task most of the time uses MFCC algorithm. There can be more than 13 MFCC features if we add 12 cepstral coefficients with 12-delta cepstral coefficient, 12 double delta cepstral coefficients, 1 energy coefficient, 1 delta energy coefficient and 1 double delta energy coefficient. Total 39 MFCC features can be calculated from a single audio file. But usually 12 MFCC features are used for audio processing. For our experiment, we only used 13 MFCC coefficient vectors. In the classification section, we have discussed that applying feature selection algorithm to the set of features, which feature is more correlated and having high coefficient value.

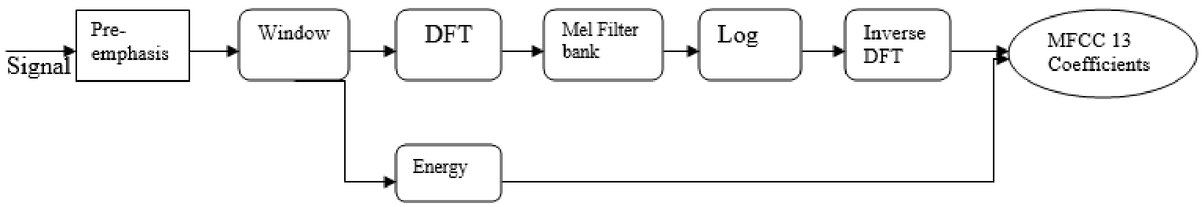


Figure 15 Steps to get MFCC

Apart from MFCC we have extracted 5 more frequency domain features. Features like Spectral Centroid, Spectral Spread, Spectral Entropy, Spectral flux and Spectral Rolloff are extracted from spectrum of audio signals. Where in spectral Centroid and Spectral Spread are not that useful for our purpose as they are usually used to count beats per minute by musicians. Spectral Spread is used by telecommunication and radio communication. For this kind of experiments Spectral Entropy is important feature which consist of the entropy of the normalized spectral energies for a set of sub-windows or sub-frames. It also describes the complexity of the signal. As appliance sound, may not be harmonic so we assumed that complexity will be there. So, to calculate spectral entropy first we must calculate spectrum $X(\omega i)$ of our signal. Along with this we need to calculate Power Spectral Density (PSD) and after that we must normalize PSD.

$$\mathbf{PSE = -\sum pi \ln pi}$$

Here, pi is the Normalized PSD and PSE is the power spectral entropy. Spectral Flux calculates that how frequently power spectrum of a signal is changing between two successive frames. Given below is the full list of the Audio features, Table 1.

Audio Features	Description
MFCCs (13) Vectors	Mel Frequency Cepstral Coefficients form a Cepstral Representation where the frequency bands are not linear but distributed per the Mel-scale.
Spectral Centroid	Midpoint of spectrum
Spectral Flux	Squared difference between spectra of two successive frames
Spectral Entropy	Entropy of the normalized spectral energies of two successive frames
Spectral Roll off	The spectral rolloff point is the portion of containers in the power range at which 85% of the power is at lower frequencies.
Spectral Spread	Technique of frequency hopping
Zero Crossing Rate (ZCR)	Rate of sign changes from positive to negative or vice versa during a particular frame
Energy	Calculate the energy of signal during a single frame
Entropy of Energy	Measure the abrupt changes in a frame

Table 1 Acoustic Features

3.3.2 Energy Features:

Energy data has been collected by Enmetric smart plug sensor and it sends the energy data to the cloud and store it there. Feature extraction is done on cloud and we can request energy data anytime. It keeps storing appliance energy information once it is plugged in. There are 8 essential features it extracts from energy signal. Resolution at which it processes energy signal is different in compare to audio signal. It has a minimum resolution of 1 minute which is big in compare to acoustic signal which is 25ms. Give below is the list of energy features, table 2:

Sr No.	Energy Features
1	Average Power (W)
2	Minimum Power (W)
3	Maximum Power (W)
4	Energy Used (W Hours)
5	Average Frequency
6	Average Voltage
7	Average Current
8	Average Power

Table 2 Energy Features

3.4 Spectral Analysis

In this section, we present some analysis we did based on spectral difference in every audio signal of appliances. When tuning into human voice, it is promptly obvious that every tune has a particular cadence or rhythm and once in a while even harmonic structure. Sounds of individual vary from those of different herds. As ahead of schedule as the eighteenth century, Barrington noticed that the tunes of cross-cultivated fowls varied from the species run of the mill melody, proposing a part for vocal learning. In any case, until the late 1950s, there had been no target method for affirming these perceptions by physical estimations of the melodies themselves. The creation of the sound spectrograph (sonogram) at Bell Laboratories was a huge achievement for quantitative examination of vocal conduct. The sonogram changes a transient a stream of sound into a straightforward static visual picture uncovering the time recurrence structure of every tune syllable. Sonogram pictures can be measured, broke down, and contrasted and each other. This permits the specialist to evaluate the level of likeness between various melodies by examining (or cross-corresponding) sonograms and arranging tune syllables into unmistakable sorts. The system is similarly valuable in looking at anomalies in appliance sounds. It is extremely normal for data to be encoded in the sinusoids that shape a signal. This is valid for normally happening signals, and those that have been made by people. Numerous things sway in our universe. For instance, discourse is a consequence of vibration of the human vocal lines; stars and planets change their shine as they turn on their tomahawks and rotate around each other; ship's propellers produce occasional removal of the water, et cetera. The state of the time domain waveform is not vital in these signs; the key data is in the frequency, phase and magnitude of the segment sinusoids. The FFT or DCT is utilized to concentrate this data.

3.4.1 Time and Amplitude

We first tried to plot time domain representation of the acoustic audio by plotting the amplitude or pressure values of audio against time. Amplitude is the fluctuation of the sinusoidal wave from its mean value. We can also say that amplitude is the maximum displacement from its equilibrium position. The difference between time domain and frequency domain is, Time domain is how the signals change over time and Frequency-domain is how much signals lie in the frequency range, as signals are composed of many sinusoidal wave signals with different frequencies. Taking an example from our experiment, we plot a graph from new microwave sound and 5-year-old microwave sound, Figure [16], [17].

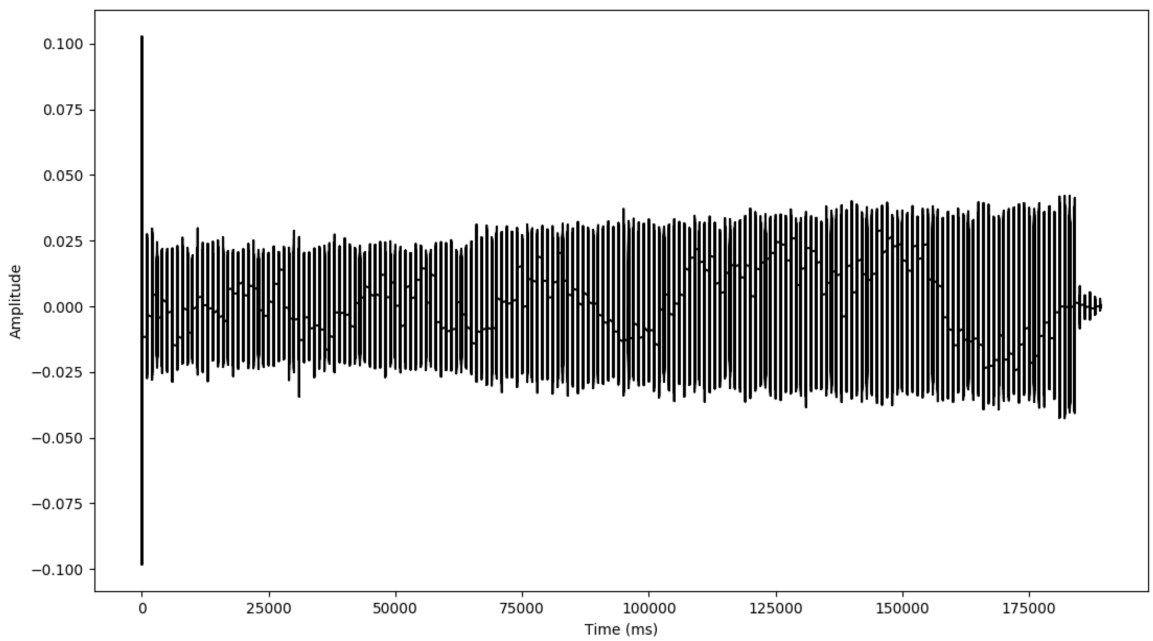


Figure 16 Amplitude and Time representation for new microwave

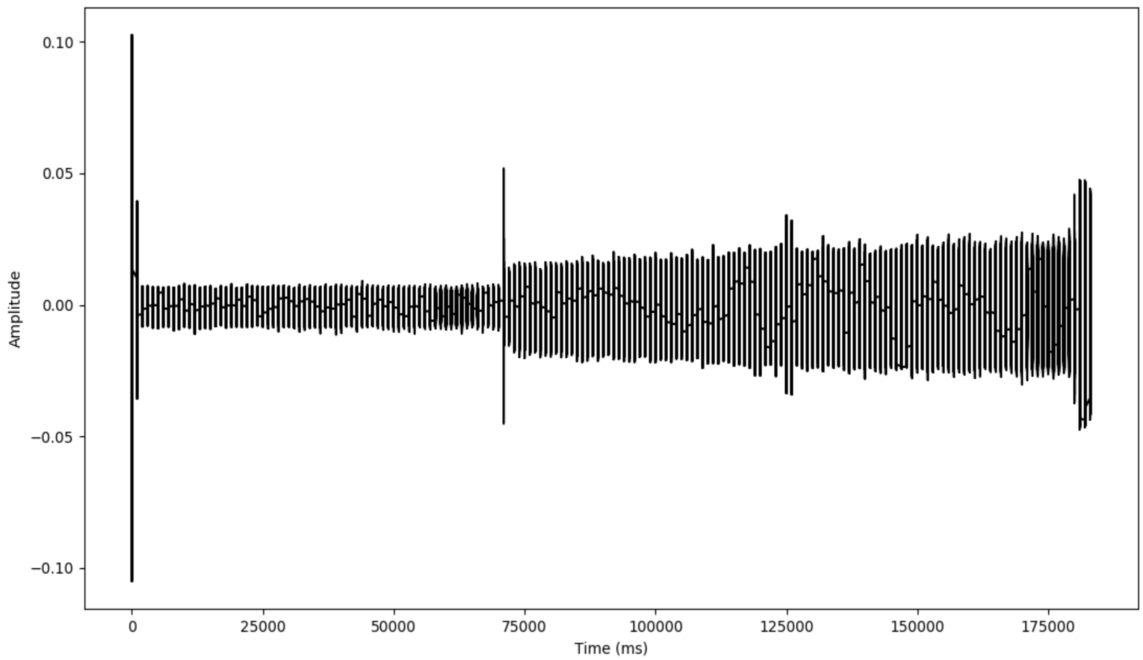


Figure 17 Amplitude and Time representation of 5-year-old microwave

In the above figures, we can observe some difference in the sound amplitude of the new and 5-year-old microwave. It is observed that 5-year-old microwave has some constant amplitude for 65000 ms or 1 minute than sudden change in amplitude and amplitude increases gradually. In figure 16, it is observed that as the time increase, an amplitude also increased gradually. We should ignore the first high amplitude spike, as that is the beep sound of the microwave when it starts. There are many pressure points in the new microwave graph. Another interpretation is that, 5-year-old microwave which is having less pressure points which means some of the parts are not working correctly and they might not be making sound. We can also look over the different microwave amplitude and time graphs.

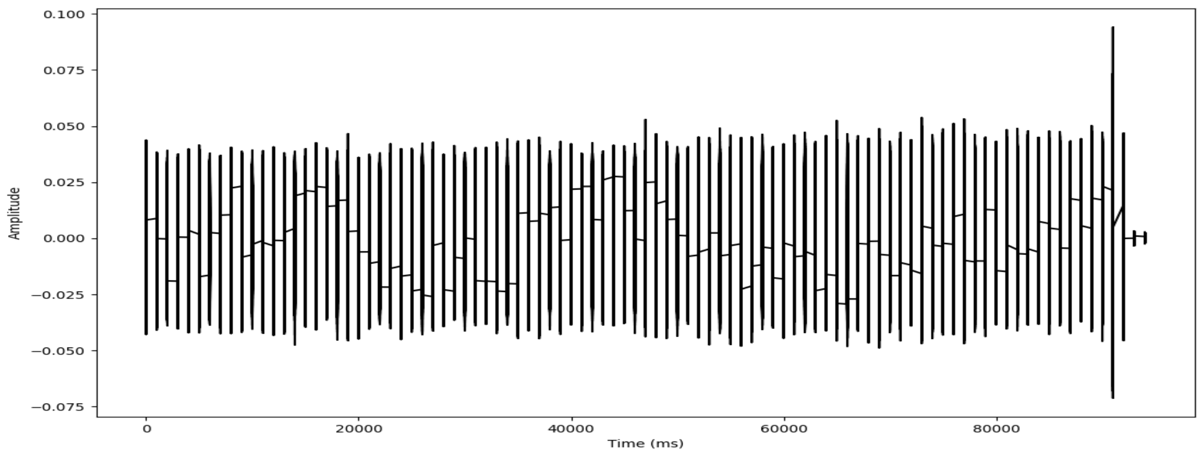


Figure 18 Amplitude and Time representation of 1-year-old microwave

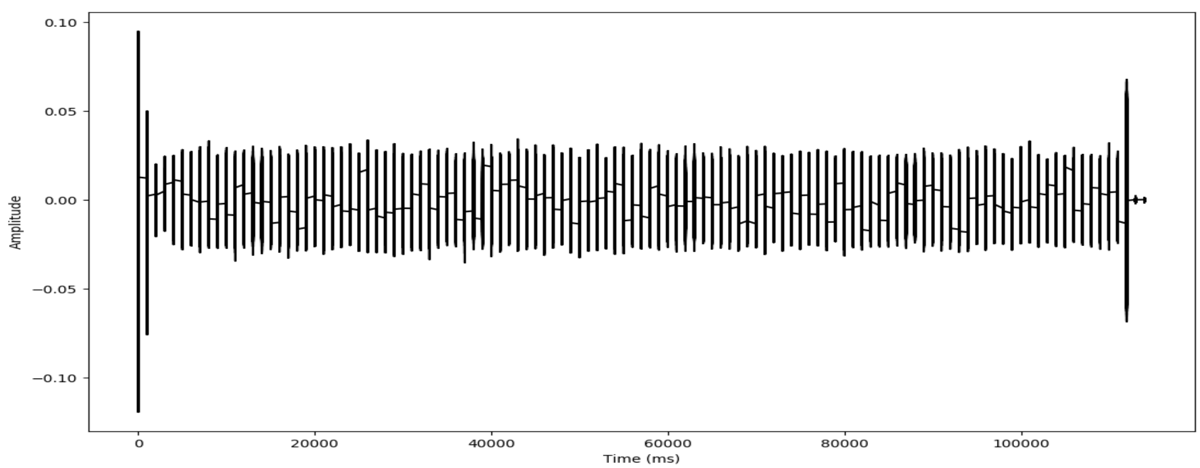


Figure 19 Amplitude and Time representation of 3-year-old microwave

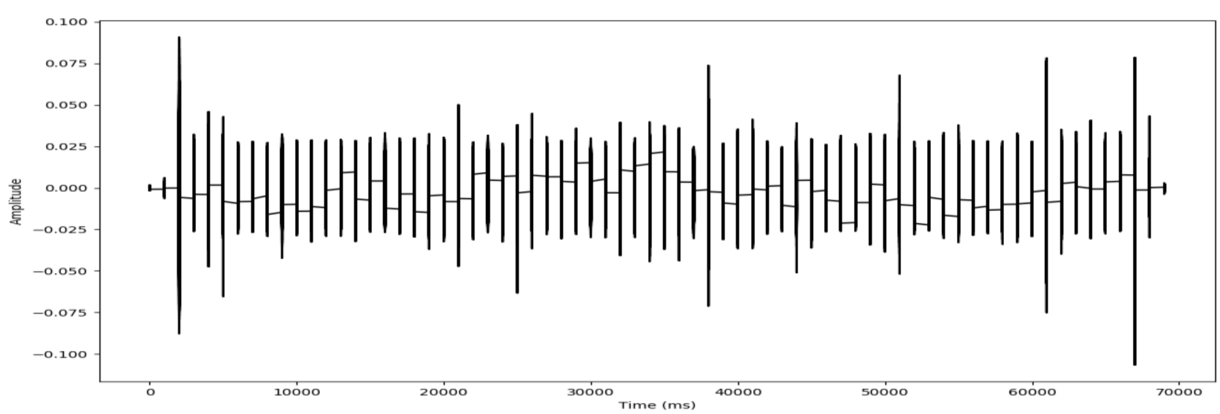


Figure 20 Amplitude and Time representation of 25-year-old microwave

3.4.2 Power Spectral Density (PSD)

To get better picture we went further to look some more pattern in data. We converted the time domain data into frequency domain data by using Fast Fourier Transform (FFT). We can inverse from frequency to time domain by applying inverse FFT. This is way to visualize power of a signal. We plotted a Power Spectral Density (PSD) graph which describes how power of signal is spread in the frequency domain. On Y-axis, there is power which is in decibels and it is calculated by taking one 10^{th} of the power of a signal. X-axis denotes frequency in Hertz.

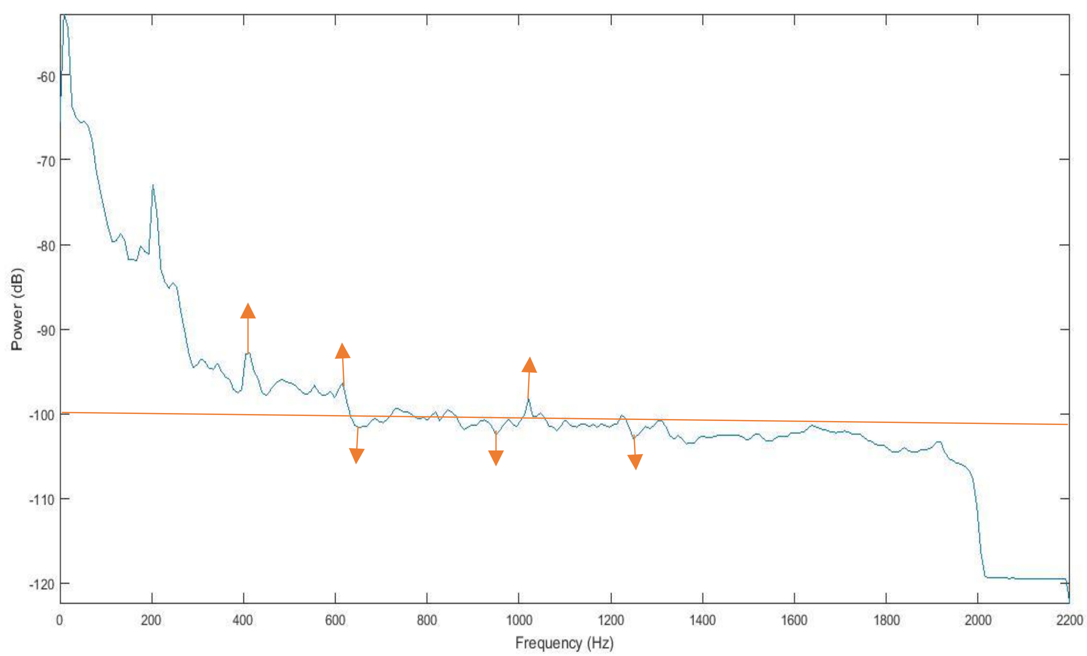


Figure 21 PSD graph for new microwave

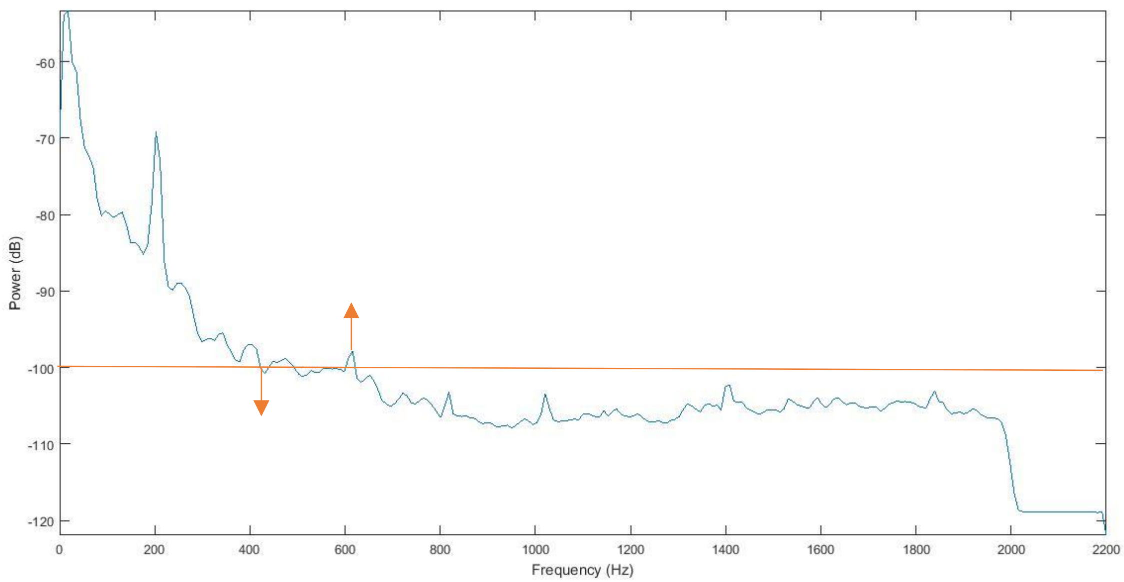


Figure 22 PSD graph for 5-year-old microwave

In above figures [21] [22], it shows the comparison between two cases, new microwave with food and 5-year-old microwave with food. It is observed that most of the frequencies in new microwave PSD are close to -100 decibels and some the frequencies are above -100 decibels. But in case of figure 22, except some lower frequencies up to 400 Hz, all other frequencies are below -100 decibels. It means higher frequencies are having lower power or intensity. Here the Power is in negative decibels. Higher the negative value, lesser the power. Above two graphs showed that there are some lower frequencies signals which is having high power. Older microwave showed some higher frequencies spikes. We also plotted Power Spectral Density graph for other acoustic sounds of microwave.

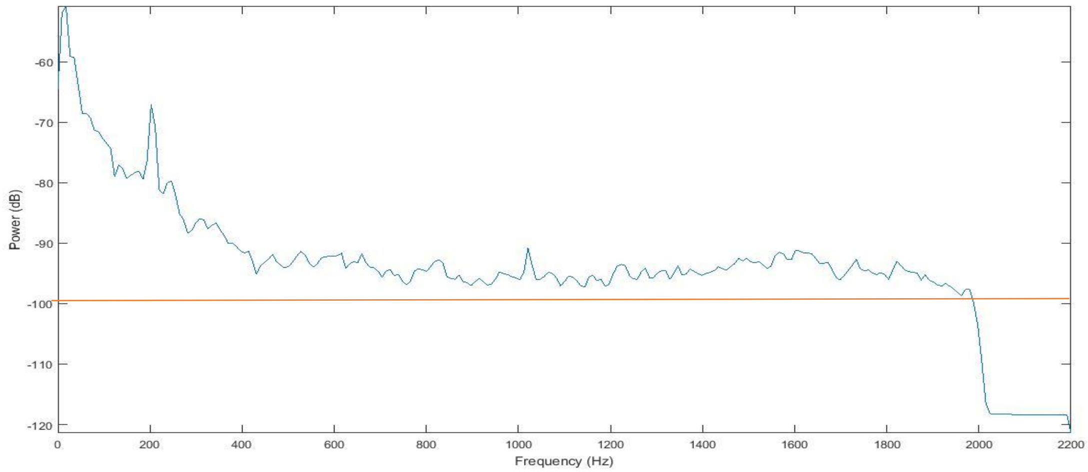


Figure 23 PSD graph for 1-year-old microwave

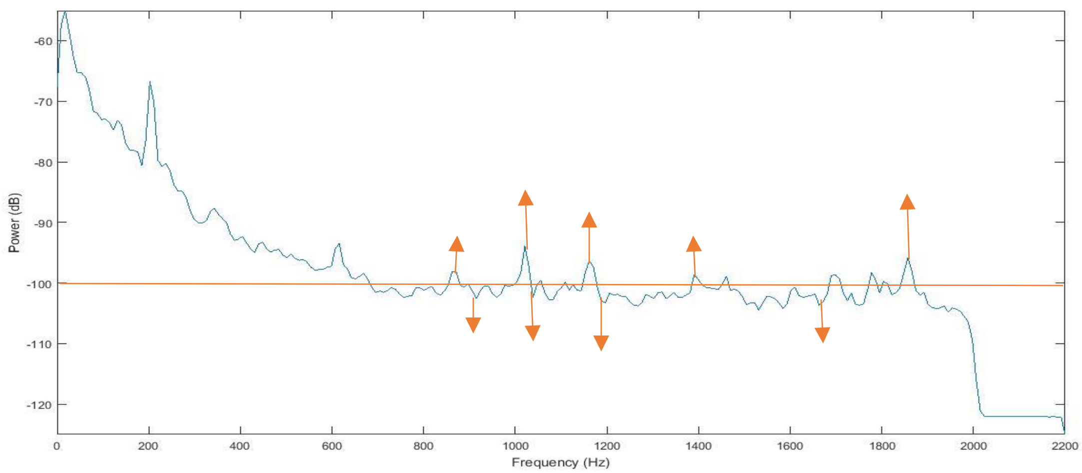


Figure 24 PSD graph for 3-year-old microwave

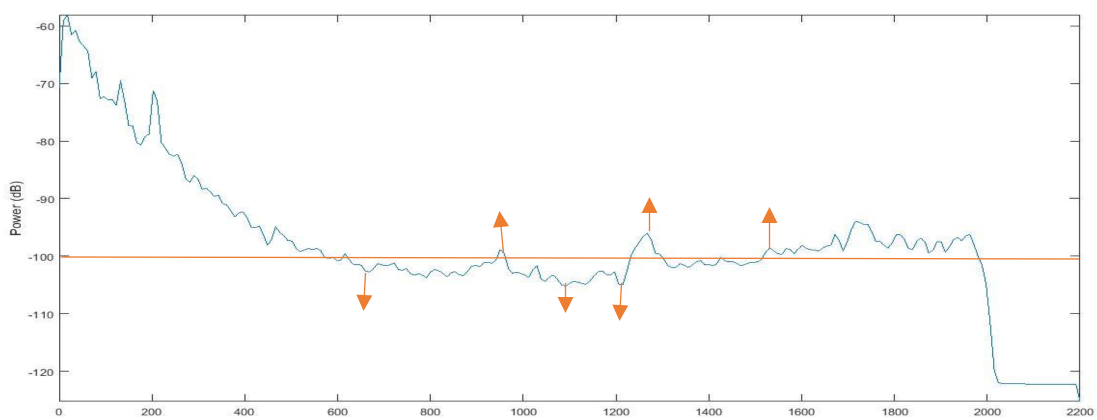


Figure 25 PSD graph for 25-year-old microwave

All graphs showed difference in frequencies with respect to power. 1-year-old microwave frequencies distributed throughout the signal with power more the -100 dB. 3-year-old microwave has less number of higher frequencies with power more than -100 dB. Oldest microwave, some fluctuation can be seen with higher frequencies. As frequencies between 600 Hz to 1600 Hz are having lower power but after that all higher frequencies are having high power more than -100 dB.

3.4.3 Time, Frequency and Power

With Power Spectral Density we are only getting power of a signal on frequency domain but at what time frequency is having what power is what we want to know. Spectrograph is an algorithm which analyses the acoustic sound at an interval of time and display time, frequency and power with a different color codes. It shows the three-dimensional view. Time is shown on the X-hub, frequency is shown on the Y-pivot and intensity is shown as the minor departure from greyscale haziness or of shading.

In below figures representing spectrograms, darker bands represent more intense components of the spectrum. Highest power is denoted with dark red and lowest with dark blue. For new microwave, light yellow color is dominant, this means that there are many frequencies between higher and lower are of medium intensity. Lower frequencies (0-2450 Hz) are of high power and highest frequencies (20000 – 22000 Hz) have lowest power. Let’s take an example of 5-year-old microwave, on a few instances there are many higher frequencies which is having high power throughout the time scale. On a 1-year-old microwave, we can see some high frequencies with high power at initial time. The dark bright yellow color is dominant in this spectrogram of 5-year-old microwave, which means most number of high frequencies are of medium power throughout the time interval.

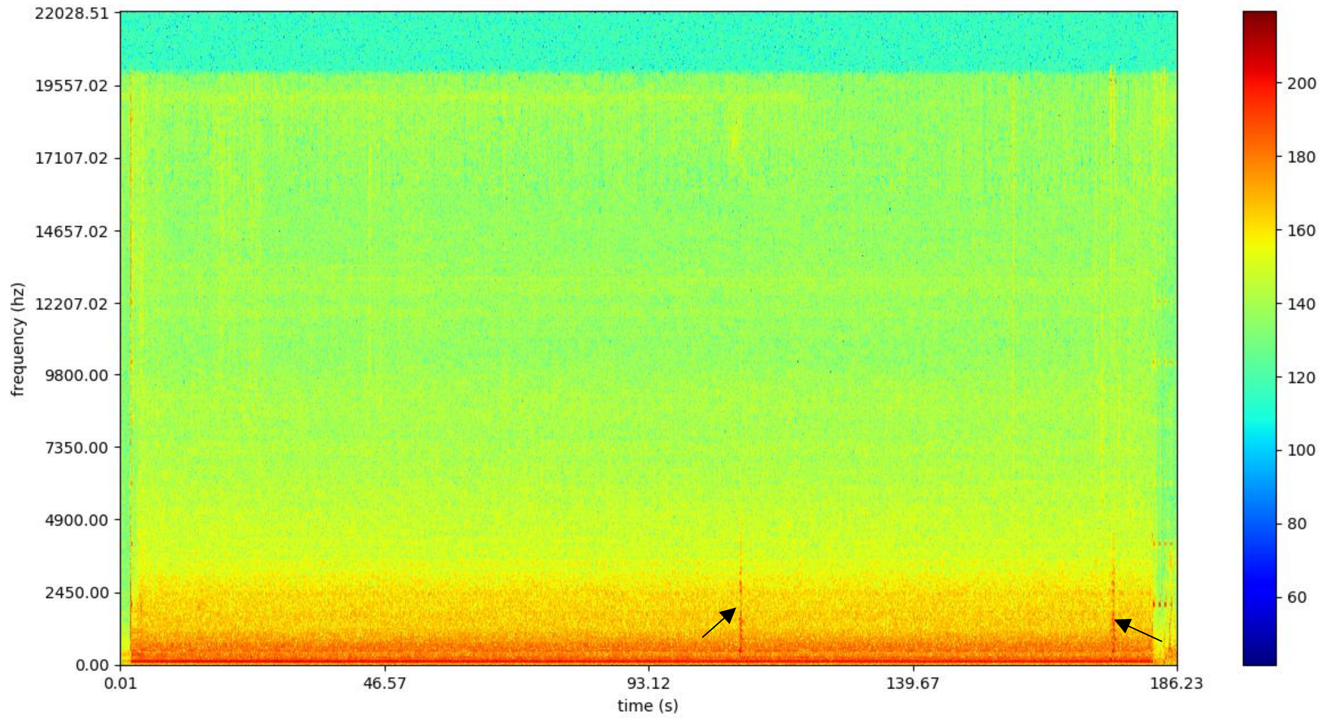


Figure 26 New Microwave Spectrogram

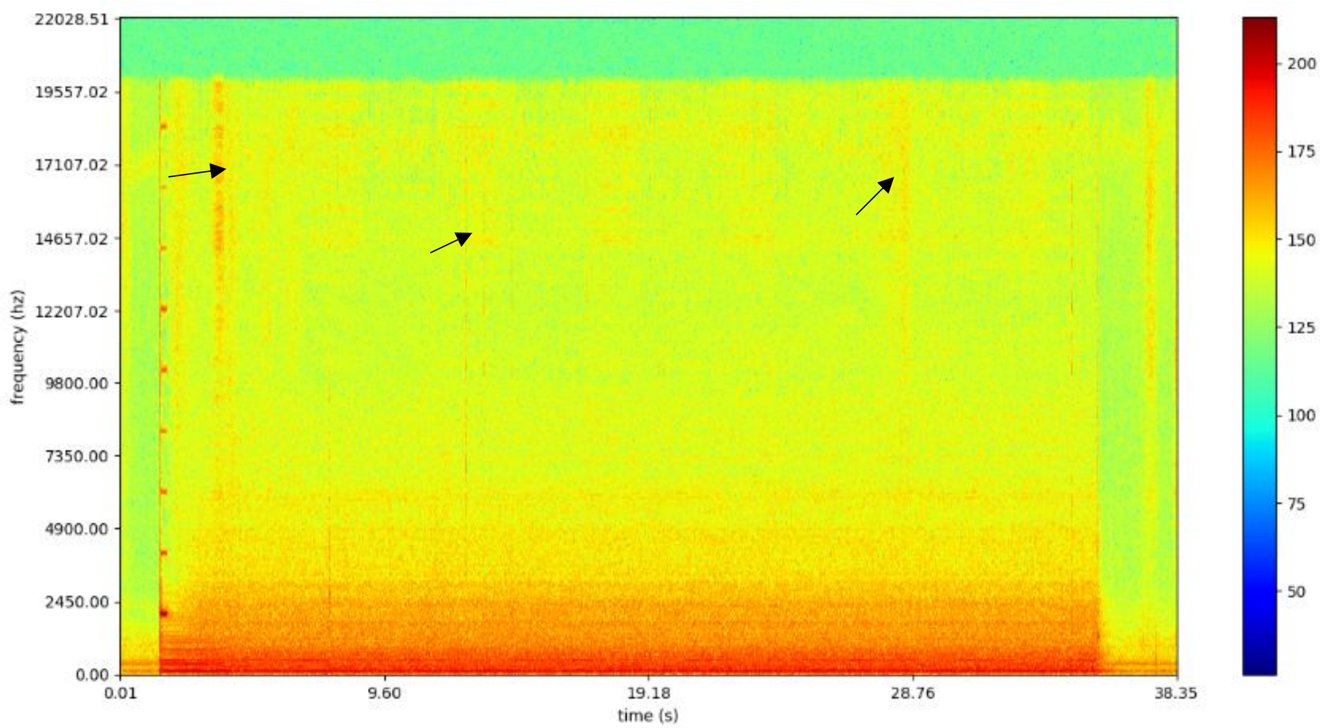


Figure 27 1- year-old Microwave Spectrogram

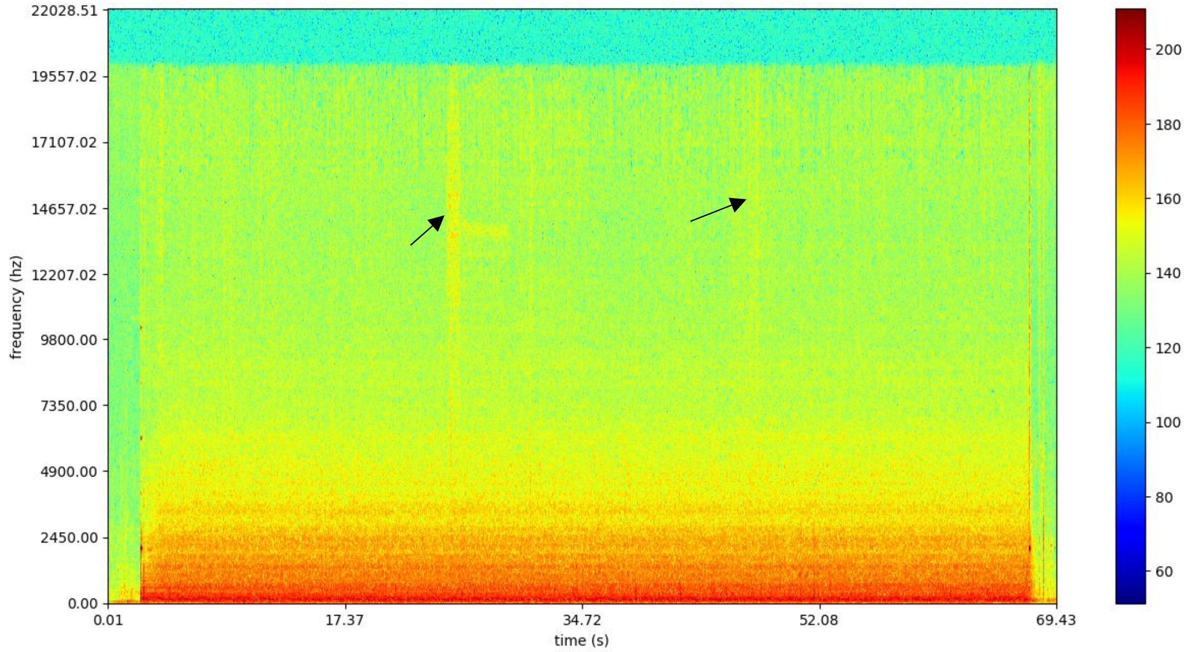


Figure 28 3- year-old Microwave Spectrogram

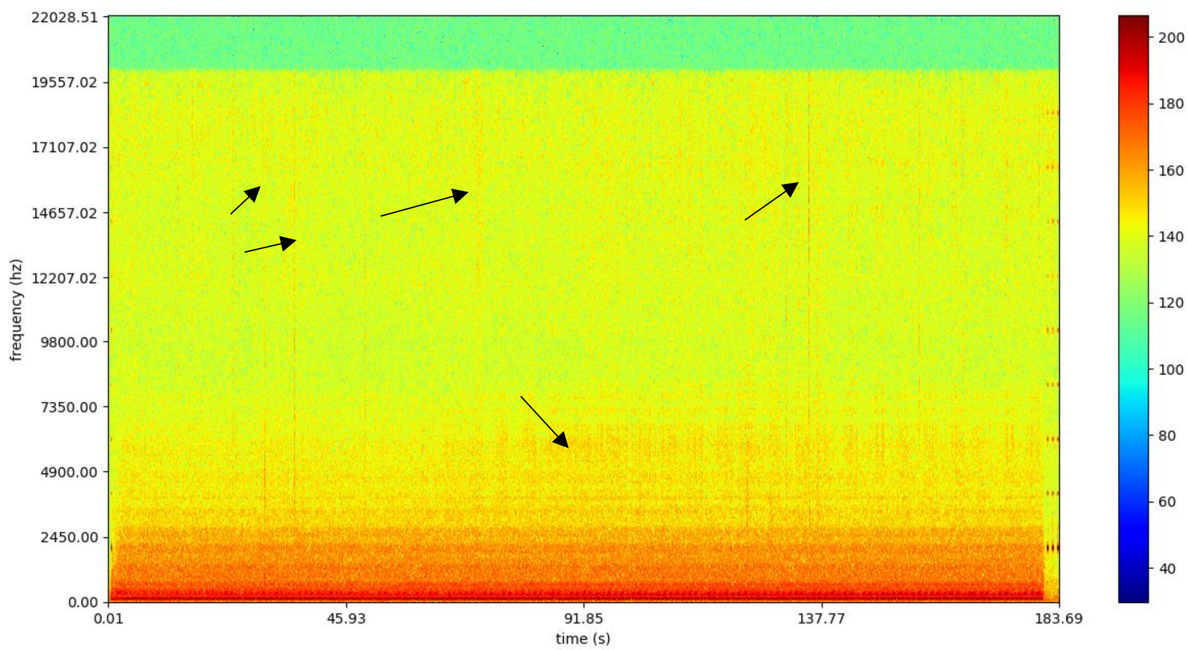


Figure 29 5- year-old Microwave Spectrogram

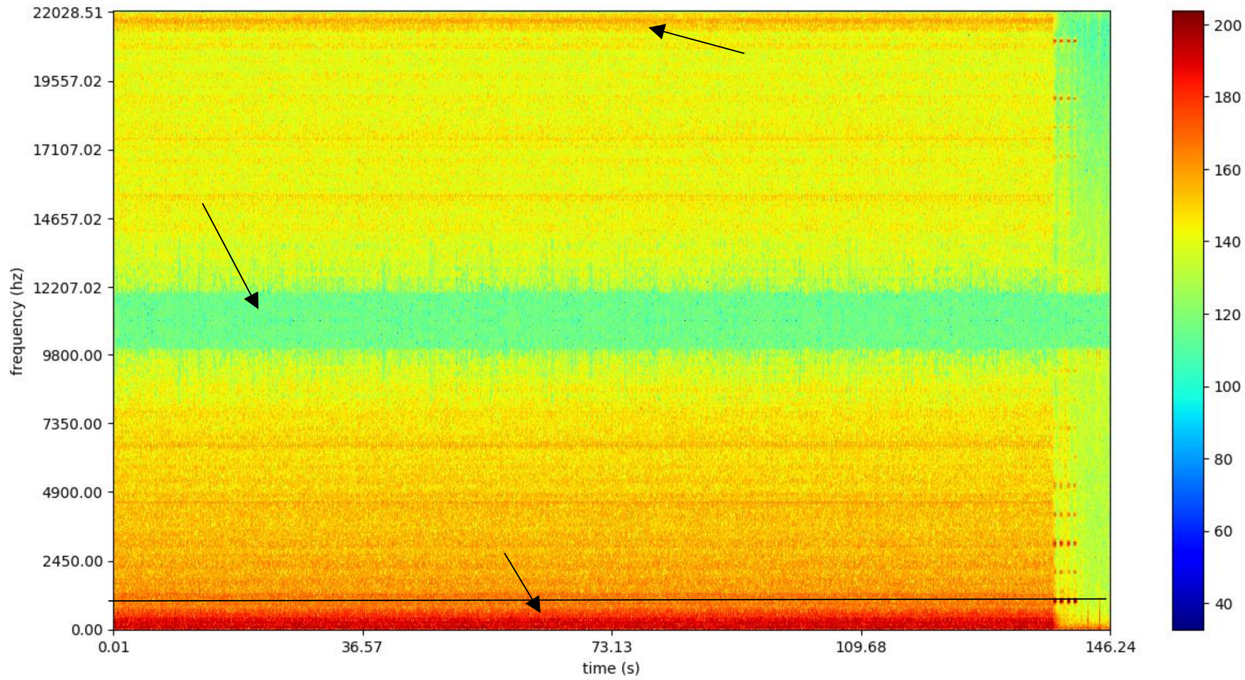


Figure 30 25- year-old Microwave Spectrogram

Every microwave spectrogram from above figures showed a difference in frequency distribution over a period with different power. There are many interpretations can be made looking over some visuals. Figure 30, We can see a drastic difference between all other microwave's spectrum and 25-year-old microwave spectrum. There are very less lower frequencies which are of high intensity throughout the time period. There are some mid frequencies between 9800 Hz to 12207 Hz which are of lower intensity. Throughout the spectrum we can see that there are alternate bands of frequencies having high and mid intensity which is denoted as red and tallow color. In all other spectrogram, we have pointed out arrow which showed the unusual frequency power over a certain time.

3.5 Classification & Data Analysis

This chapter focuses on the classification of acoustic sound and energy data. Section 3.5.1 focuses on preparation of data set which includes labeling which is the first step towards supervised learning. We also give a brief description of different test cases we performed before jumping into the detecting anomalies. Test cases such as classifying a microwave with food or without food. Classifying old microwave with food and new microwave with food. We performed all these test cases based on acoustic features and energy features separately. It also explains why we have choose supervised learning and its advantages. Section 3.5.2 explains acoustic data classification which includes feature selection algorithm and shows accuracy result on different algorithm. Also, explains the confusion matrix and gives the ROC graph. Section 3.5.3 explains energy data classification with different algorithms applied on dataset. Section 3.5.4 explains the hybrid model where classification algorithm was used on the combined hybrid dataset containing acoustic and energy features. Some data analysis techniques are used like pruning, boosting, oversampling to get good accuracy.

3.5.1 Preprocessing

Our aim is to detect whether appliance state is normal or abnormal. To do this we labeled each feature file manually with a class of normal and abnormal. But before directly jumping into detecting anomaly, we tried to classify different use cases. First, we tried to test our system to classify a microwave with food and without food. So, we took new microwave acoustic and energy data of with and without food. Extract feature and label them as withFood and withoutFood. In WEKA, first we run machine learning algorithm (Support Vector Machine) for acoustic data.

3.5.1.1 Test Case 1

First case we tried to classify a microwave with food and without food from acoustic feature set. This test we performed to check that how accurate is our acoustic features. Figure 31, is the final dataset with binary class "withFood" and "withoutFood".

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V
Zero Cros	Energy	Entropy o	Spectral C	Spectral S	Sprectral I	Spectral F	Spectral R	Mfcc1	2	3	4	5	6	7	8	9	10	11	12	13	Class
0.07078	0	1.477897	0.384712	0.244496	2.435956	0	0.483666	-42.9456	0.75	-0.02	0.25	0.2084	-0.0099	0.332112	-0.35963	0.039115	-0.38327	0.358829	0.16268	-0.43683	withoutFo
0.100272	0.008901	1.032527	0.136587	0.174295	0.350257	0.005523	0.050817	-23.3853	2.24	0.071	0.33	0.0627	0.13804	0.025534	0.051188	0.004452	0.030737	0.010172	0.023635	0.008543	withoutFo
0.033122	0.00903	1.13802	0.135078	0.17115	0.346205	0.000032	0.049909	-23.3871	2.23	0.069	0.33	0.0608	0.13625	0.023721	0.049377	0.002537	0.028713	0.008028	0.021362	0.006544	withoutFo
0.009982	0.000237	2.240662	0.136993	0.196387	0.21002	0.002188	0.026316	-29.3016	2.31	0.138	0.38	0.0578	0.16718	0.07309	0.112235	0.075489	0.107641	0.059281	0.051422	0.034256	withoutFo
0.060345	0.000108	0.624945	0.152345	0.19626	0.377877	0.002773	0.053539	-29.3721	2.29	-0.01	0.3	0.0107	0.09192	0.004047	0.043023	0.012614	0.039871	-0.00814	-0.00286	-0.01735	withoutFo
0.110254	0.000001	2.572085	0.329882	0.316547	0.720565	0.003507	0.058984	-38.1391	2.34	-0.06	0.26	0.0277	-0.066	0.294503	0.329197	0.353146	0.386907	0.223157	0.060216	-0.06844	withoutFo
0.122505	0	3.251564	0.319618	0.319558	0.620496	0.003107	0.043557	-39.4285	2.36	-0.48	0.34	0.0659	-0.2983	-0.09221	0.250574	0.392633	0.400935	0.210769	0.035739	-0.13426	withoutFo
0.165154	0	3.250888	0.341227	0.323125	0.747127	0.000962	0.049002	-39.3859	2.47	-0.5	0.36	0.3032	-0.3387	-0.37459	-0.04197	0.27219	0.374567	0.133977	-0.09852	0.036162	withoutFo
0.1951	0	3.271713	0.346824	0.318615	0.997958	0.000733	0.155172	-39.0746	2.29	-0.51	0.11	0.3675	-0.2438	-0.03654	-0.00272	0.165207	0.527311	0.245852	-0.06767	0.009299	withoutFo
0.170599	0	3.25952	0.341226	0.303898	1.114572	0.000569	0.179673	-39.0755	2.03	-0.19	0.2	0.2586	-0.345	-0.03331	-0.04199	0.123889	0.252833	0.28135	0.133997	-0.00899	withoutFo
0.171506	0	3.247971	0.34302	0.300068	1.253129	0.000711	0.23775	-39.1044	1.9	-0.11	0.12	0.1381	-0.4009	-0.2195	-0.13597	-0.09441	0.054503	0.079488	-0.07746	-0.0006	withoutFo
0.181488	0	3.255221	0.340421	0.303123	1.184951	0.000561	0.218693	-39.0619	1.92	-0.35	-0.01	0.1066	-0.4283	-0.17971	-0.00893	-0.18434	-0.01625	-0.12642	-0.00674	0.044687	withoutFo
0.163793	0	3.282444	0.328484	0.304154	1.061457	0.000554	0.157895	-38.8201	2	-0.2	0.05	0.3696	-0.2098	-0.11556	0.110929	-0.06523	0.159093	-0.02592	0.065452	-0.10772	withoutFo
0.144283	0	3.269669	0.322405	0.303649	0.906633	0.000675	0.098004	-38.8664	2	-0.07	0.32	0.2766	-0.0816	-0.08902	0.062005	-0.05522	0.091088	0.058452	0.289898	-0.11492	withoutFo
0.115699	0	3.278897	0.318782	0.305326	0.776958	0.000807	0.061706	-38.9211	2.04	-0.24	0.37	0.1716	-0.1144	-0.02628	0.105697	0.132881	0.154213	0.088747	0.082862	0.004451	withoutFo

C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V
1.548319	0.1712	0.224991	0.3771	0.005514	0.051724	-23.4424	2.236073	0.055545	0.325844	0.084428	0.130994	0.024942	0.060352	-0.00895	0.040184	0.005654	0.027887	-0.00482	withFood
1.726979	0.169572	0.223066	0.368004	0.000079	0.050817	-23.4483	2.230773	0.049973	0.320078	0.078733	0.12547	0.019803	0.055243	-0.01408	0.034993	0.000604	0.0229	-0.01033	withFood
2.353772	0.15124	0.218603	0.187997	0.003619	0.021779	-29.1829	2.301069	0.154344	0.392459	0.136623	0.187978	0.068909	0.093212	0.056685	0.090007	0.050867	0.057662	0.020802	withFood
0.862155	0.155711	0.199331	0.385255	0.004261	0.053539	-29.3922	2.222328	0.082397	0.324008	0.074494	0.141143	0.040828	0.07929	0.02898	0.024441	-0.02603	0.024612	0.020522	withFood
2.06228	0.241523	0.300708	0.230994	0.007102	0.014519	-38.0709	2.467639	0.179688	0.311047	0.070797	0.207548	0.099431	0.126537	0.220681	0.241274	-0.03134	-0.01188	0.131024	withFood
3.257099	0.328641	0.318551	0.796832	0.006279	0.04265	-39.4287	2.556904	0.163588	0.103639	0.072121	0.010132	0.0069	0.45773	0.265868	0.123254	0.040549	0.153589	0.002173	withFood
3.302699	0.320654	0.306767	0.987645	0.001209	0.143376	-38.3098	1.997212	-0.13922	0.318252	0.012342	0.069409	0.087518	0.216659	0.253372	0.252986	0.189921	0.189732	-0.071	withFood
3.250214	0.31918	0.305677	1.049087	0.000831	0.176044	-38.1058	1.938277	-0.1779	0.277912	-0.33179	-0.05118	0.169039	0.284645	0.074192	0.138155	0.036508	-0.00503	0.064057	withFood
3.174305	0.301481	0.307353	0.738554	0.0011	0.059891	-38.4466	2.423677	0.060823	0.234337	-0.41521	-0.17912	0.325453	0.365624	0.01147	0.205265	0.052761	0.087689	0.095092	withFood
3.097764	0.316753	0.305825	0.968044	0.000999	0.142468	-38.6295	2.363555	0.059106	0.002155	-0.29304	-0.06638	0.226284	0.44039	0.160961	0.102237	-0.00902	0.139765	-0.01565	withFood
3.163196	0.33205	0.305518	1.104451	0.00078	0.196915	-38.9102	2.223247	-0.00923	0.103896	-0.27133	-0.16295	0.141257	0.569996	0.313186	0.07093	-0.00053	0.122656	0.088994	withFood
3.262973	0.343069	0.314576	0.929298	0.000837	0.088929	-39.3055	2.575071	-0.10388	0.264745	-0.185	-0.05498	0.200149	0.42307	0.424448	0.060559	-0.0172	0.095803	0.216235	withFood
3.244271	0.358544	0.315116	0.973396	0.001245	0.18784	-39.6125	2.561614	-0.01973	0.212784	-0.19791	-0.01761	0.094547	0.375115	0.355652	0.057744	-0.02767	0.049757	-0.01257	withFood

Figure 31 Acoustic feature set of case with and without food

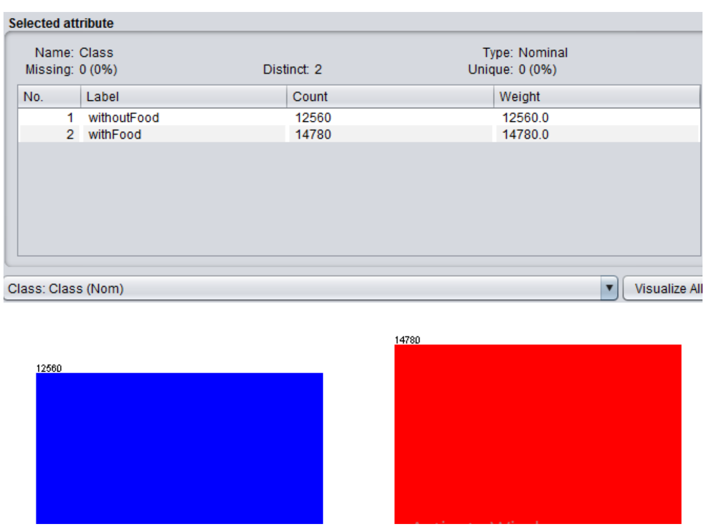


Figure 32 Class Distribution for Test Case1

The above figure 32, explains the two classes and withFood class instances are more than withoutFood. With all features, we checked accuracy by running SVM algorithm to this dataset with 10 folds cross-validation.

```
=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances       25351               92.7249 %
Incorrectly Classified Instances      1989                7.2751 %
Kappa statistic                          0.8539
Mean absolute error                      0.0728
Root mean squared error                  0.2697
Relative absolute error                 14.6467 %
Root relative squared error             54.1233 %
Total Number of Instances            27340

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.939    0.082    0.906      0.939    0.922      0.854    0.928     0.879     withoutFood
                 0.918    0.061    0.946      0.918    0.932      0.854    0.928     0.913     withFood
Weighted Avg.    0.927    0.071    0.928      0.927    0.927      0.854    0.928     0.897

=== Confusion Matrix ===

     a     b   <-- classified as
 11790   770 |     a = withoutFood
  1219 13561 |     b = withFood
```

Figure 33 Test Case1 confusion matrix

This is the first test case accuracy we have got approx. 93%. True positive rate is also high around 93% and false positive rate is low as 8.2% with a precision of 90%. We want to check which ifeature is more useful as we have used MFCC's 13 vectors and other features. So, to check which feature is more useful and can give us more better accuracy. So, we run feature selection algorithm based on correlation coefficient value we ranked the features. Figure 34 and 35, displayed the ranking of features based on the correlation coefficient value and information gain.

```
        Correlation Ranking Filter
Ranked attributes:
 0.3134    18 10
 0.2846    14 6
 0.2461    15 7
 0.232      2 Energy
 0.214     21 13
 0.2058    16 8
 0.2047    17 9
 0.1543     7 Spectral Flux
 0.1296     5 Spectral Spread
 0.1284     3 Entropy of Energy
 0.1187    19 11
 0.0764    11 3
 0.0714    13 5
 0.0681    12 4
 0.0625    20 12
 0.0526     9 Mfcc1
 0.0509     4 Spectral Centroid
 0.0505     1 Zero Crossing Rate
 0.0474    10 2
 0.0432     6 Sprectral Entropy
 0.0252     8 Spectral RollOff
```

Figure 34 Correlation Ranking

```
Ranked attributes:
 0.40028     9 Mfcc1
 0.39792     2 Energy
 0.28746     6 Sprectral Entropy
 0.16151     8 Spectral RollOff
 0.1236      1 Zero Crossing Rate
 0.10419    14 6
 0.07992     3 Entropy of Energy
 0.07799    18 10
 0.06707    15 7
 0.05976    12 4
 0.05628     5 Spectral Spread
 0.05324     7 Spectral Flux
 0.0377      4 Spectral Centroid
 0.03759    17 9
 0.0371     21 13
 0.03505    16 8
 0.02558    10 2
 0.01753    11 3
 0.01546    19 11
 0.0066     13 5
 0.00375    20 12
```

Figure 35 Information Gain Ranking

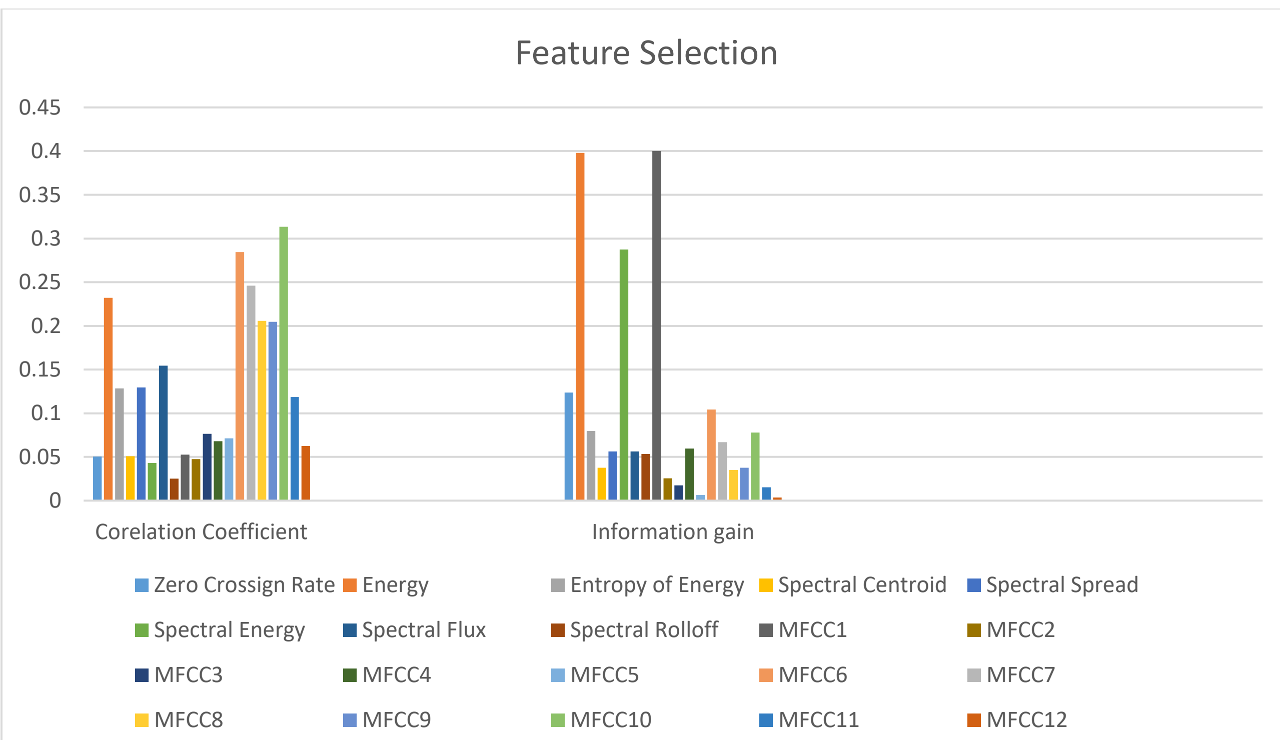


Figure 36 Feature selection comparison for Test Case1

Above figure 36, is the feature selection algorithm results based on two criteria, one is correlation coefficient and other is information gain. We selected top 5 features from each result and run the learning algorithm to check the accuracy of the model. Top five features that we have selected from correlation coefficient ranking are MFCC, Energy, Spectral Flux, Spectral Spread, Entropy of Energy. From information gain, selected attributes are MFCC, Energy, Spectral Entropy, Spectral RollOff, Zero Crossing Rate.

```
=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances       20522               75.0622 %
Incorrectly Classified Instances      6818               24.9378 %
Kappa statistic                          0.4913
Mean absolute error                      0.2494
Root mean squared error                  0.4994
Relative absolute error                 50.2066 %
Root relative squared error            100.2065 %
Total Number of Instances            27340

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.642    0.157    0.777      0.642    0.703      0.498    0.742     0.663     withoutFood
                 0.843    0.358    0.735      0.843    0.785      0.498    0.742     0.704     withFood
Weighted Avg.    0.751    0.266    0.754      0.751    0.747      0.498    0.742     0.685

=== Confusion Matrix ===

     a     b   <-- classified as
  8060  4500 |     a = withoutFood
  2318 12462 |     b = withFood
```

Figure 37 Confusion Matrix of features selected from correlation coefficient

```
=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances       20929               76.5508 %
Incorrectly Classified Instances      6411               23.4492 %
Kappa statistic                          0.5256
Mean absolute error                      0.2345
Root mean squared error                  0.4842
Relative absolute error                 47.2096 %
Root relative squared error             97.1695 %
Total Number of Instances            27340

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.712    0.189    0.762      0.712    0.736      0.527    0.761     0.675     withoutFood
                 0.811    0.288    0.768      0.811    0.789      0.527    0.761     0.725     withFood
Weighted Avg.    0.766    0.243    0.765      0.766    0.765      0.527    0.761     0.702

=== Confusion Matrix ===

     a     b   <-- classified as
  8940  3620 |     a = withoutFood
  2791 11989 |     b = withFood
```

Figure 38 Confusion Matrix of features selected from information gain

We received 75% accuracy with top 5 features from correlation ranking. Accuracy was decreased from full set of features. It is incorrectly classifying instances, as it can be seen from confusion matrix. Features selected per information gain gives better accuracy about 76%. With this much of information, we can still conclude that our model can classify two classes of acoustic features. We tried same techniques for energy feature dataset. We classify this dataset by applying J48 Decision tree algorithm.

3.5.1.2 Test Case 2

In this test case, we have tried our classification model with energy feature set. We have performed the same techniques and methods that we did for acoustic data in test case1. In this test case, we again want to classify a microwave from with food and without food.

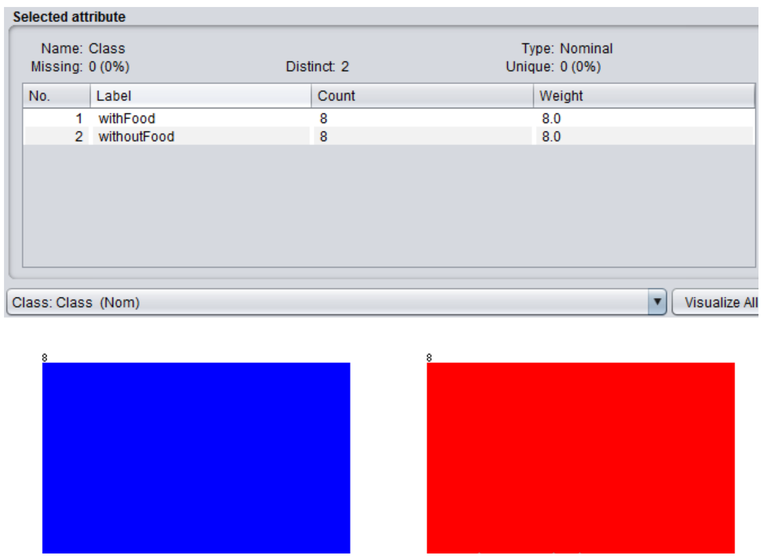


Figure 39 Energy data class distribution

By applying algorithm to all the features, we got 75% accuracy with 62.5% true positive rate and 12.5% false positive rate. Following figure 40, shows the result of our experiment with energy feature set.

```
=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances          12               75      %
Incorrectly Classified Instances         4               25      %
Kappa statistic                          0.5
Mean absolute error                      0.3125
Root mean squared error                  0.4977
Relative absolute error                 61.5942 %
Root relative squared error             98.0684 %
Total Number of Instances               16

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.625    0.125    0.833      0.625    0.714      0.516    0.719     0.675     withFood
                 0.875    0.375    0.700      0.875    0.778      0.516    0.719     0.680     withoutFood
Weighted Avg.    0.750    0.250    0.767      0.750    0.746      0.516    0.719     0.677

=== Confusion Matrix ===

 a b   <-- classified as
 5 3 | a = withFood
 1 7 | b = withoutFood
```

Figure 40 Test case2 classification result

We applied j48 algorithm and checked which feature is taking into consideration. To check this, we visualize a tree.

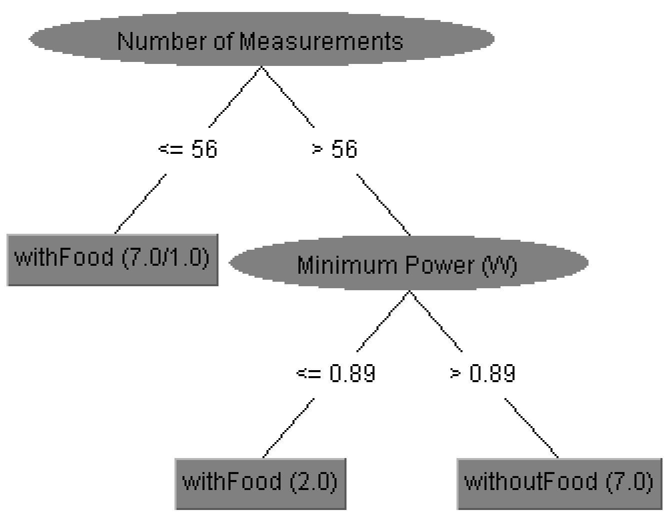


Figure 41 Test case 2 tree visualization

To improve accuracy, we tried feature selection algorithm. According to correlation coefficient we have selected top 5 features. But accuracy didn't change from 75% when we applied j48. We applied different algorithm in top 5 features to improve accuracy.

```
=== Attribute Selection on all input data ===

Search Method:
	Attribute ranking.

Attribute Evaluator (supervised, Class (nominal): 10 Class ):
	Correlation Ranking Filter
Ranked attributes:
 0.20851   5 Average Frequency
 0.12451   9 Number of Measurements
 0.07607   1 Average Power (W)
 0.07607   4 Energy Used (W Hours)
 0.07208   7 Average Current
 0.05966   6 Average Voltage
 0.0216    8 Average Power Factor
 0.00778   3 Maximum Power (W)
 0.00111   2 Minimum Power (W)

Selected attributes: 5,9,1,4,7,6,8,3,2 : 9
```

Figure 42 Test case 2 energy feature selection result

We train and tested this data set with different algorithms. Figure 43, shows the accuracy of different algorithm on this dataset. Among all OneR performed best to classify a microwave with food or without food.

Algorithm	**Accuracy**
SVM	43.75%
Naïve Bayes	43.75%
IBK	43.75%
Boosting	68.75%
Bagging	81.25%
J48	75%
OneR	81.25%

Table 3 Test case 2 classification algorithm accuracy comparison

3.5.1.3 Test Case 3

Now, the last we tried is the combined the acoustic and energy features, which is our hybrid classifier. We want to investigate the accuracy of this hybrid model and we will see which features are important to classify between with food and without food. It is difficult to merge these two features because resolution is different for both the sensors. Energy features has 1 minute of resolution and acoustic has 0.25ms of resolution. So, in preprocessing we took average of 1 minute of data instances of acoustic for every one minute of energy data. So, that we can match the values between them. But later, while extracting features we set a frame or window of 1 minute instead of 25 ms. Now we have equal number instances of acoustic data and energy data.

7	8	9	10	11	12	13	Average P	Minimum	Maximum	Energy Us	Average F	Average V	Average C	Average P	Number o	CLASS
0.120908	0.293927	0.182917	0.219401	0.007672	0.12783	0.212208	1.04	0.91	1.17	0.017333	60.2	118.72	0	0	60	withoutFood
0.142269	0.365262	0.204495	0.259802	0.023414	0.187125	0.179532	576.54	0.94	1042.82	9.609	60.2	116.62	5.733	0.8	59	withoutFood
0.195426	0.370643	0.150066	0.20167	-0.03958	0.134244	0.11278	1015.39	995.36	1029.17	16.92317	60.1	115.06	9.685	0.9	59	withoutFood
0.243233	0.370205	0.114526	0.161934	-0.04416	0.134392	0.134184	978.72	959.81	999.07	16.312	60.1	114.85	9.248	0.9	60	withoutFood
0.207414	0.342639	0.066701	0.140266	0.017547	0.114964	0.145642	405.67	1.15	966.04	6.761167	60.2	117.13	3.835	0.8	60	withoutFood
0.250517	0.225712	-0.02837	0.348489	0.198915	0.204234	0.123081	850.21	1.25	962.93	14.17017	60.1	115.87	8.087	0.9	57	withoutFood
0.297949	0.287937	-0.03096	0.327148	0.19892	0.208598	0.156089	910.79	896.03	922.74	15.17983	60.2	115.85	8.41	0.9	59	withoutFood
0.160773	0.308456	0.084678	0.135496	0.090826	0.216766	0.066553	1.11	0.85	1.22	0.0185	60.1	118.31	0	0	60	withFood
0.154574	0.260165	0.012479	0.070061	0.000461	0.192578	0.028798	718.89	1.21	972.57	11.9815	60.2	115.62	7.02	0.8	56	withFood
0.171212	0.256942	0.000313	0.032343	-0.04002	0.177536	0.026805	972.78	967.25	979.26	16.213	60.1	114.77	9.197	0.9	56	withFood
0.144558	0.246526	-0.01563	0.010317	-0.04128	0.148151	0.044382	974.26	970.06	979.42	16.23767	60.1	114.74	9.212	0.9	56	withFood
0.198695	0.287321	0.020491	0.216596	0.037042	0.145929	0.150207	311.7	0.88	980.05	5.195	60.2	117.14	3.089	0.8	57	withFood
0.075038	0.204925	-0.01033	0.253481	0.072065	0.188664	0.12313	978.52	974.1	983.37	16.30867	60	114.76	9.276	0.9	55	withFood
0.049932	0.195833	-0.00028	0.244536	0.112138	0.215839	0.136117	982.58	974.79	986.7	16.37633	60.1	115.24	9.28	0.9	56	withFood

Figure 43 Hybrid feature set of test case 3

Above figure 44, is the hybrid dataset having acoustic and energy features. Followed by same techniques used in test case 1 and test case 2 we tested our model with different classification algorithms. The result we get are listed in the following table.

Classification Algorithm	Accuracy
SVM	71.428%
NaïveBayes	71.426%
J48	50%
OneR	42.857%
IBK	92.85%

Table 4 Test case 3 accuracy comparison

3.5.1.4 Test Case 4

Now, we experimented another case where we tried to classify between new and old microwave. We followed same techniques to check the accuracy of our model. We have used new microwave and 5-year-old microwave. Labeled them manually as new and old. We performed each classification separately for acoustic, energy and hybrid.

First, we took acoustic feature set and run classification algorithm on it. Table 5, shows the result of accuracy.

Classification Algorithm	**Accuracy**
SVM	99.57%
NaiveBayes	98.67%
J48	99.36%
OneR	98.13%

Table 5 Acoustic features accuracy to classify into old and new microwave

Second, we took energy feature set and run classification algorithm on it. Table 6, shows the result of accuracy.

Classification Algorithm	**Accuracy**
SVM	81.25%
NaiveBayes	18.75%
J48	62.5%
OneR	56.25%
IBK	75%

Table 6 Energy features accuracy to classify into old and new microwave

Third, we inspected hybrid model by running classification algorithm on hybrid data set. Table 7, shows the accuracy.

Classification Algorithm	**Accuracy**
SVM	65.21%
NaiveBayes	95.65%
IBK	93.06%
OneR	94.45%
J48	91.30%

Table 7 Table 6 Hybrid features accuracy to classify into old and new microwave

To improve the accuracy, we used feature selection algorithm and experiments shows that combination of features such as MFCC's, Average Frequency, Average Voltage and Number of Measurements gives the highest accuracy on each learning algorithm. When we run SVM on full feature set, it has given 65.21% accuracy. But after feature selection its giving 78.26% accuracy with very low false positive rate around 22%, figure 44.

```
=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances          18               78.2609 %
Incorrectly Classified Instances         5               21.7391 %
Kappa statistic                          0.5525
Mean absolute error                      0.2174
Root mean squared error                  0.4663
Relative absolute error                 45.0579 %
Root relative squared error             94.7325 %
Total Number of Instances               23

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.786    0.222    0.846      0.786    0.815      0.555    0.782     0.795     NEW
                 0.778    0.214    0.700      0.778    0.737      0.555    0.782     0.631     old
Weighted Avg.    0.783    0.219    0.789      0.783    0.784      0.555    0.782     0.731

=== Confusion Matrix ===

  a  b   <-- classified as
 11  3 |  a = NEW
  2  7 |  b = old
```

Figure 44 Test case 4, summary of classifier on selected features

After features been extracted from acoustic and energy data of all microwaves, labeling is the next step. As it is part of pre-processing where all microwaves feature dataset is merged and we manually labeled them into two classes:

Normal: We have labeled normal to new, 1-year-ol, 3-year-old and 5-year-old microwaves. This is supervised learning and we would expect good accuracy based on all our previous test cases. Labeling of acoustic feature set and energy feature set is done separately. So, that later we can perform separate classification for acoustic and energy.

Abnormal: We have manually labeled faulty microwave as an abnormal. As we have less instances of abnormal classes, so we have performed some oversampling. Before oversampling, we also checked the accuracy of the model. In later section, we have compared different accuracies.

3.5.2 Acoustic Data Classification

Once labeling is complete for acoustic feature set, we run the feature selection algorithm on whole acoustic feature set to check which feature is more helpful. Following table will explains the accuracy of different feature combinations on different classification algorithm. Initially we run complete feature set on 5 classification algorithms (SVM, Naïve Bayes, J48, OneR, IBK). We checked the accuracy of the model. Among all, J48 gave the highest accuracy about 99.02%. After that we used feature selection algorithm to check the ranking of all acoustic features. Based on correlation coefficient, we have selected top five features i.e. MFCC, Spectral Spread, Spectral Flux, Zero Crossing Rate and Energy. With these top 5 features we again run the classification algorithm on it to check the accuracy. Apparently, accuracy decreased for J48. But it is not decreased too much. Still J48 is giving highest accuracy. After some iterations with different combination of features, we have used on MFCC and checked the accuracy. Support Vector Machine is giving best outcome. SVM shows consistent good results on all cases. It gives accuracy of 98.63% on MFCC, 98.55% on Full set and 98.59% on top 5 features.

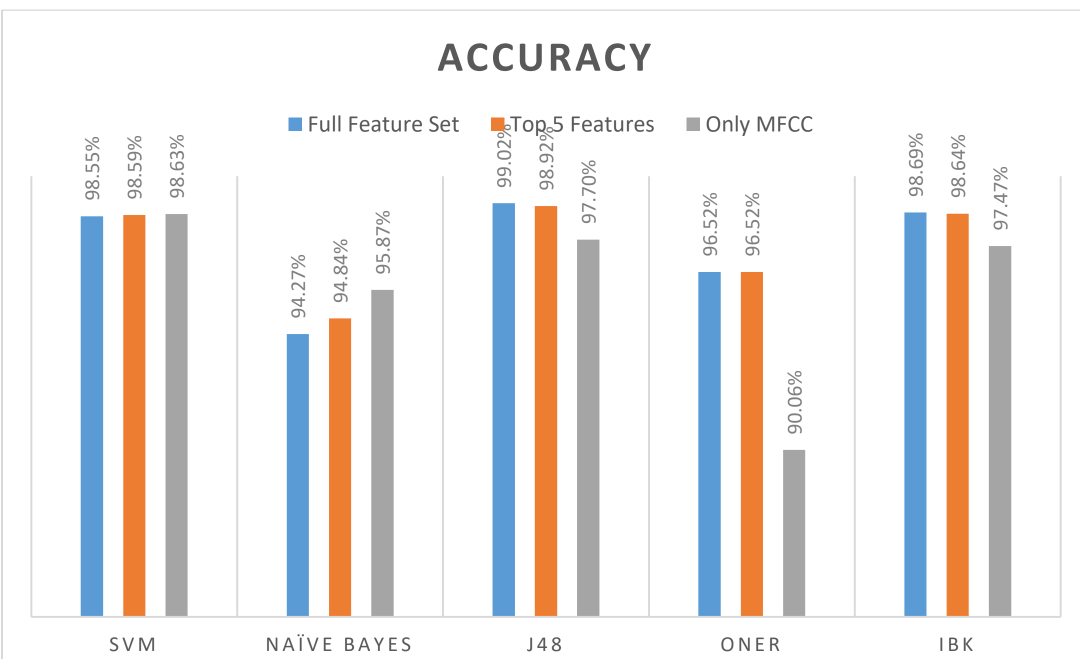


Figure 45 Acoustic classifier accuracy comparison

3.5.3 Energy Data Classification

The energy classifier classifies the microwave into normal and abnormal using the energy feature data. Feature file downloaded from Enmetric cloud is labeled manually into normal and abnormal. Dataset is split into training and testing set. We have used several classification algorithms to train and test the model. Cross validation with 10-Folds applied. First, we recorded the accuracy of the model with the full feature set. We also looked upon the confusion matrix to see false positive rate and true positive rate. Following table 8, shows the accuracy of a model when different algorithm was applied.

Classification Algorithm	Accuracy
Support Vector Machine	64.10%
Naïve Bayes	53.84%
J48	46.15%
IBK	53.84%
Boosting	66.66%

Table 8 Energy classifier accuracy on all energy features

Boosting algorithm is giving highest accuracy among all. As we observed that false positive rate is high for all the algorithms, which is not a good result for this context. To lower this false positive rate and to get better accuracy we tried feature selection algorithm. Feature selected based on correlation coefficient and information gain. We have selected top five features based on their ranking on each attribute evaluator.

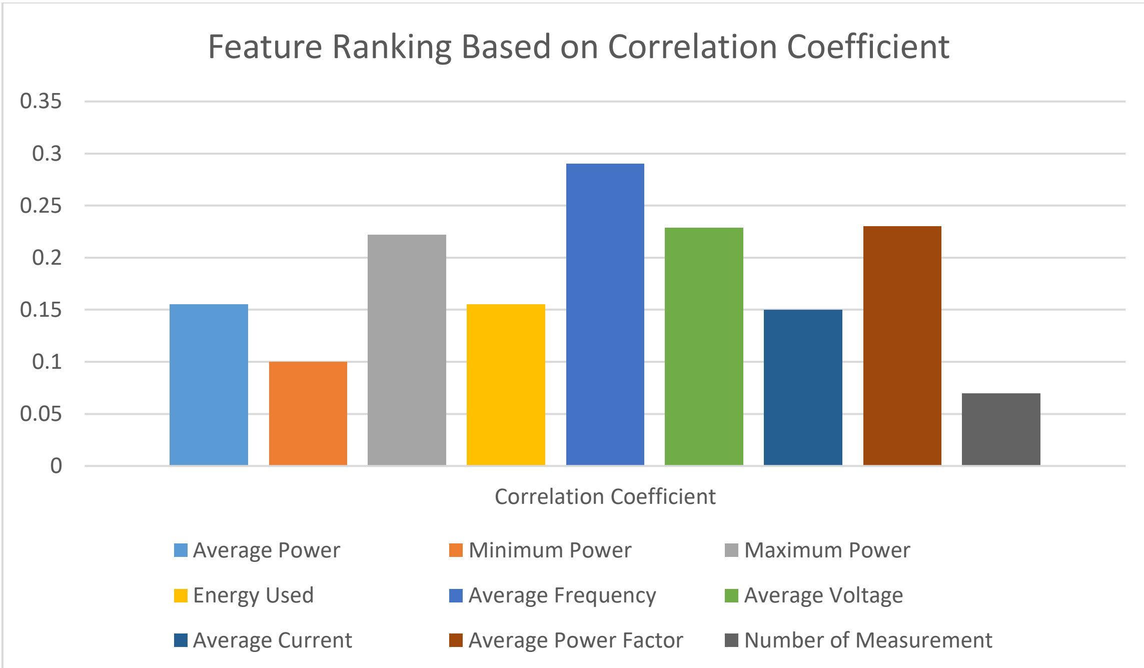


Figure 46 Energy features ranking based on correlation coefficient

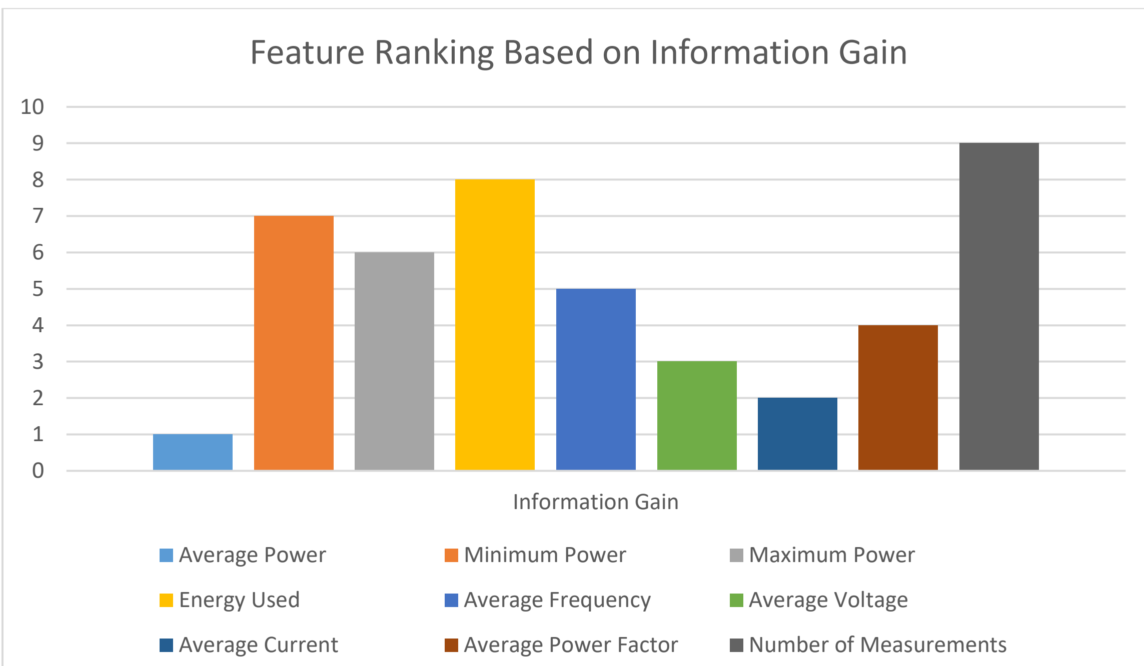


Figure 47 Energy features ranking based on information gain

From correlation coefficient, we have selected top 5 features based on their rank i.e. Average Frequency, Average Power Factor, Average Voltage, Maximum Power, Average Current. From Information gain, top five features are Number of Measurements, Energy Used, Minimum Power, Maximum Power, Average Frequency. With this top features, we tried to run classification algorithm to check the accuracy. Following table shows the result.

Classification Algorithm	**Accuracy**
Support Vector Machine	71.79%
Naïve Bayes	56.41%
J48	53.84%
IBK	48.71%
Boosting	74.35%

Table 9 Energy selective features classification accuracy based on correlation coefficient

With information gain based feature ranking, we also selected top 5 features and run machine learning algorithm on it. Below table with show the comparison between different learning algorithms on selective features based on information gain.

Classification Algorithm	**Accuracy**
Support Vector Machine	64.10%
Naïve Bayes	53.84%
J48	58.97%
IBK	53.84%
Boosting	69.23%

Table 10 Energy selective features classification accuracy based on information gain

Based on above observations, it proved that features selected from correlation coefficient ranking gives better accuracy.

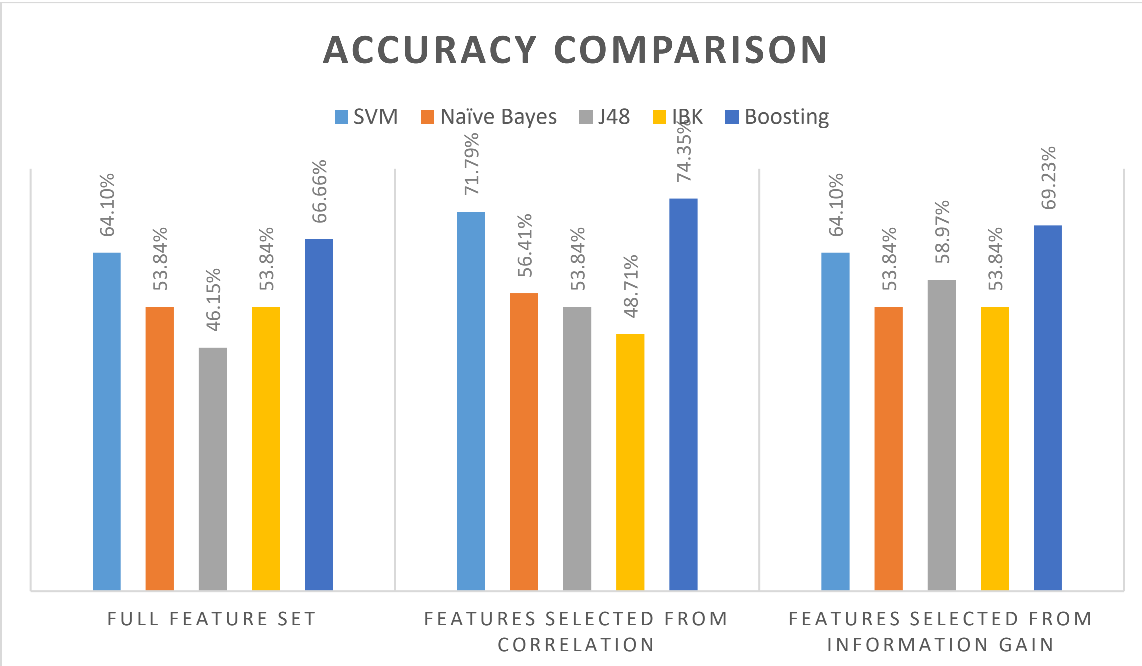


Figure 48 Energy classifier accuracy comparison

This means that energy features like Average Frequency, Average Power Factor, Average Voltage, Maximum Power, Average Current are very useful to classify any appliance condition is normal or abnormal. Training and testing model with support vector machine and boosting algorithm shows good accuracy on these features. As the abnormal microwave instances are less, which makes this dataset skewed. That is why boosting algorithm works good with this kind of dataset as boosting machine learning algorithm converts the weak learners into strong ones. Boosting algorithm is used with decision tree which is usually called as best out of box classifier.

3.5.4 Hybrid Data Classifier

Our thesis brings very novel idea to combine energy and acoustic data to detect anomaly in an appliance. In above discussion, we have seen the accuracy of our acoustic classifier engine and energy classifier engine. This is hybrid classifier engine which classifies the hybrid data which is the combination of energy and acoustic data. Hybrid dataset has all the features from acoustic and energy feature set. In preprocessing step, we changed the resolution of acoustic data so that all instances will have same resolution and overall number of instances will remain same. We labeled hybrid dataset normal for all the microwave from new to 5-year-old and abnormal for the faulty microwave. We have applied multiple classification algorithms to a full set of hybrid features to check whether our model can classify normal and abnormal or not.

Classification Algorithm	Accuracy
Support Vector Machine	58.06%
Naïve Bayes	90.32%
OneR	91.02%
IBK	87.66%
Boosting	94.32%

Table 11 Hybrid classifier accuracy comparison on whole feature set

After getting accuracy to a full set of features we ranked features by running feature selection algorithm. We rank features based on their correlation coefficient value. The features that we have selected are MFCC, Maximum Power, Average Frequency, Average Volume, Average Current and Average Power Factor. Combination of these features gave us highest accuracy and very low false positive rate in one of our classification algorithm. Accuracy of some classification algorithm decreased from these feature set.

Classification Algorithm	Accuracy
Support Vector Machine	60.83%
Naïve Bayes	93.54%
OneR	91.32%
IBK	90.32%
Boosting	94.09%

Table 12 Hybrid classifier accuracy comparison on selective feature set

Chapter 4: Comparison of The Results

This chapter demonstrate the different level of accuracies, confusion matrix, F-measure, ROC for acoustic, energy and hybrid-data classifier. As we have run machine learning algorithm on each feature set and recorded accuracy of that. To summarize, we created a comparison chart between all classifier engine (Acoustic, energy and hybrid-data) based on the algorithm which gave best accuracy. We repeated same comparison for the top ranked feature set. This comparison also determines that how accurately a classifier engine is classifying the normal and abnormal instances.

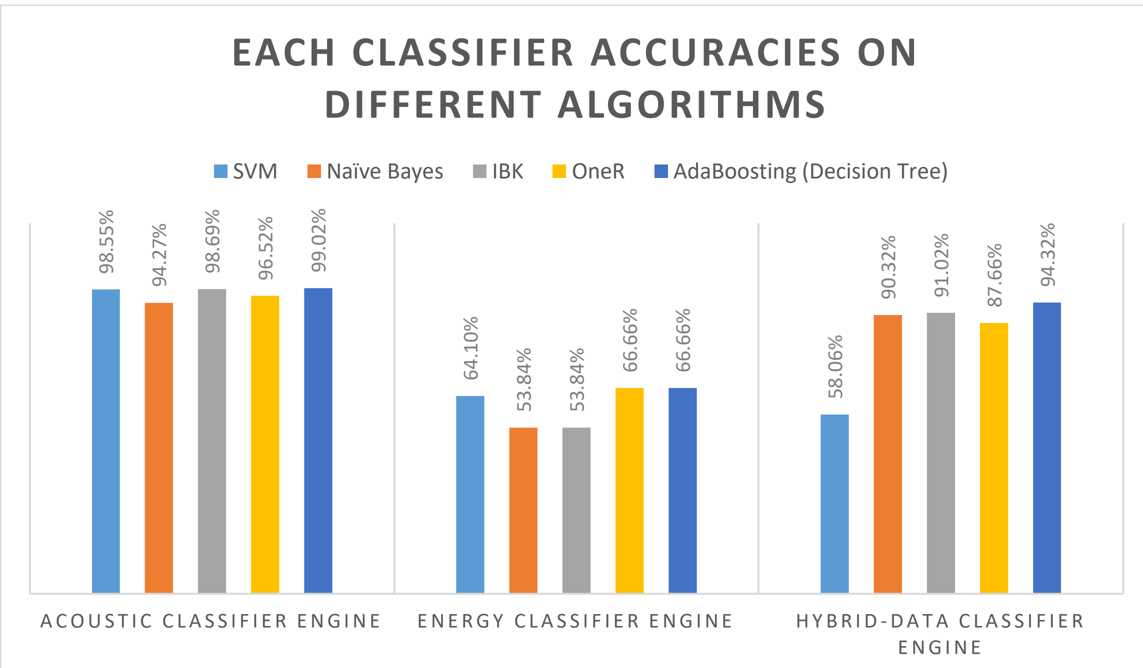


Figure 49 Each classifier accuracy on different algorithms

We also performed experiments by selecting features based on their correlation coefficient rank. Summary of all the accuracies are given below:

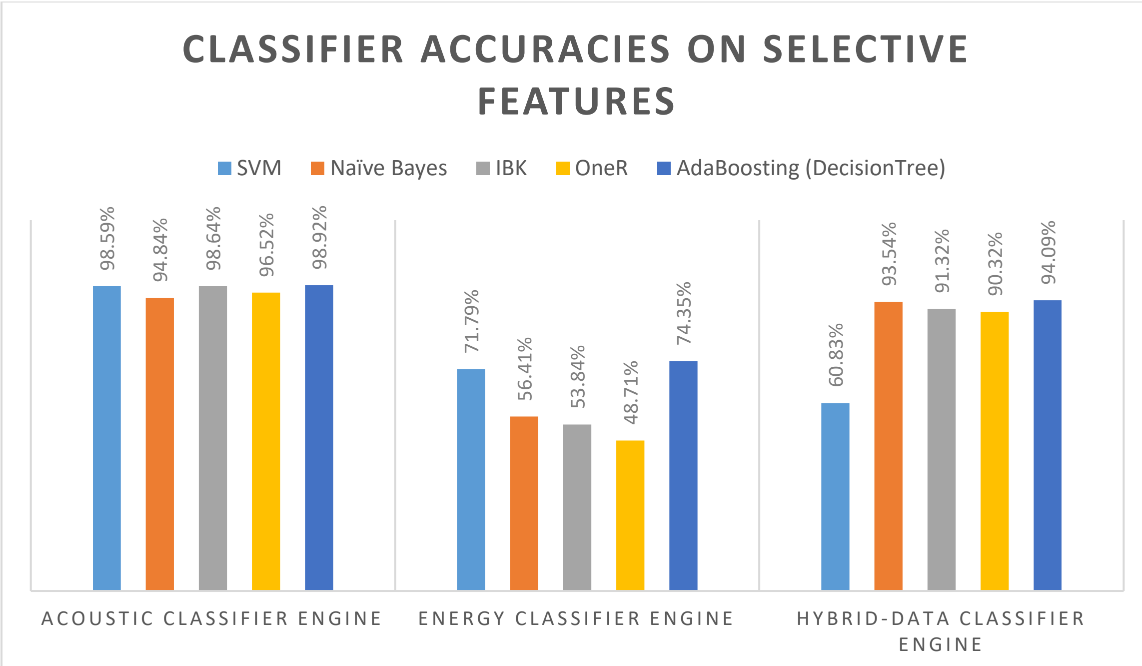


Figure 50 Classifier accuracies on selective features based on correlation coefficient

Another comparison we did is F-measure which is weighted harmonic mean of the precision and recall. We took f-measure of all highest accuracy provider algorithm from each classifier. With the same information, we also compared precision, recall, true positive rate and false positive rate.

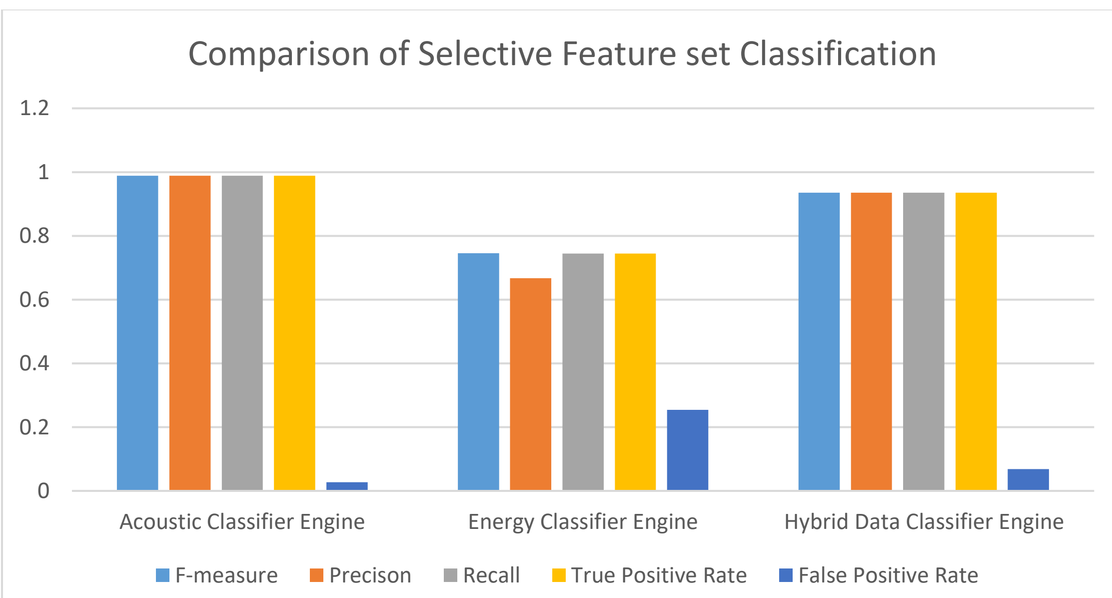


Figure 51 Summary of best learning algorithm on each classifier

Chapter 5: Conclusion

This thesis introduces novel approach for detecting anomaly in smart home appliances by creating a hybrid data from acoustic sensor and energy sensor. Our experiments proved that the accuracy of an acoustic classifier engine is very high, it means anomaly detection through acoustic sound is possible and it can be achieved with a higher accuracy. We also evaluated that our model can distinguish between different age of appliances, which can be very helpful for manufacturer in making warranty or guarantee policy of their products. We have performed several experiments on acoustic data, energy data and hybrid data with different machine learning algorithms applied to train and test the model. We found that support vector machine and decision tree algorithm works good enough to give high accuracy in classification on every dataset. We also performed feature selection algorithm on every data set i.e. acoustic, energy and hybrid. For acoustic feature set top 5 features which gives high accuracy are MFCC, Spectral Flux, Spectral Spread, Zero Crossing Rate, Energy. Among all this features we also found that MFCC is the strongest feature for acoustic which itself can give good accuracy of 98.63%. We found that frequency domain features are more informative for acoustic type of data set. As time domain feature only gave 80% accuracy with high false positive rate. For energy dataset, features like Average Frequency, Average Power Factor, Average Voltage, Maximum Power and Average Current are among top ranked features. We also infer that, for energy dataset, feature extraction algorithm based on correlation coefficient is better than information gain. We also showed some spectral analysis of acoustic data, from which we infer that some unusual events cannot be observed are clearly can visualized and it can be shown as time, frequency and amplitude graph. We found that the darker color shows the intensity of the frequency at a given instances. By looking to the spectrogram of new and other microwaves such as faulty or old, many things can be inferred such as if new one is showing some high frequencies with high intensity and old microwave is not showing any high frequencies with high intensity, this means any machinery part of old microwave may not be working. By looking at new and faulty microwave, we distinguish some difference. We also proved that using less sensor data, anomaly can be detected.

Chapter 6: Limitations and Future Work

For appliance health estimation, we focused on analyzing anomaly in an appliance by combing different sensor data without including any operating context. The biggest challenge for this kind of work is to combine two different sensor data, as both sensor we used have a different resolution. We did not know how frequently an appliance is been used. For example, a new microwave can be used for 12-15 times a day. But someone will use new microwave once a week. That is why it is necessary to include the context to get approximation. We could not have collected different appliance data as it is very difficult to find appliance of different time and an appliance which is flawed. A vibration sensor can also be incorporated with current project. It will provide the x,y,z axis detail. Any appliance which gets old consume lot of energy and starts getting vibration. Sometimes it is observed that an appliance also gets heat up quickly. So, to get the appliance body temperature is also one of the thing can be taken into consideration for detecting an anomaly. It very necessary how appliance operates and how it fails. This information is missing in our current project which can be proposed for future work. We only worked on to investigate that whether it is possible to detect an anomaly in appliances using less sensor data when no context information is given. In future, devices like onion omega2, raspberry pi can be used to collect acoustic data and compute it within and send it to server through wifi. It will save time and increase an efficiency. Automatic feedback generation could be done so that customer should know and that feedback will be send to manufacturer so that they can ask come for a service. There is still lot of work should be done in collecting acoustic data and extracting human voice from appliance noise.

References

[1] Bong Jun Ko, Jorge Ortiz, Theodoros Salonidis, Maroun Touma, Dinesh Verma. (2016). Acoustic Signal Processing for Anomaly Detection in Machine Room Environments. 2.

[2] Charlie Mydlarz, S. N. (2014). The design of urban sound monitoring devices. *Audio Engineering Society*.

[3] Charlie Mydlarz, S. N. (2014). The implementation of MEMS microphones for urban sound sensing. *Audio Engineering Society*, 9.

[4] Enmetric. (n.d.). *ENMETRIC SYSTEMS SMART PLUG*. Retrieved from ENMETRIC SYSTEMS: https://www.enmetric.com/

[5] Goodwill. (n.d.). *Chesapeake inc*. Retrieved from http://www.goodwillches.org/

[6] Jaime Lloret, E. M. (2012). Ubiquitous Monitoring of Electrical Household Appliances. *Sensors 2012, 12, 15159-15191; doi:10.3390/s121115159*.

[7] Kiran K. Rachuri, M. M. (2010). A Mobile Phones based Adaptive Platform for Experimental Social Psychology Research. *UbiComp '10, Sep 26–29, 2010, Copenhagen, Denmark*, 281-290.

[8] Leonardo Badino, C. C. (May 2014). An auto-encoder based approach to unsupervised learning of subword units. *ICASSP, IEEE International Conference on Acoustics, Speech and Signal.*

[9] Linxia Liao, T. H. (n.d.). Device Health Estimation by Combining Contextual Control Information with Sensor Data and Device Health Prognostics Utilizing Restricted Boltzmann Machine. 11.

[10] MD Abdullah Al Hafiz Khan, S. L. (2015). A Microphone Sensor based System for Green Building Applications. *IEEE International Conference on Pervasive Computing and Communications Demonstrations*, 3.

[11] Meseguer, N. A. (2009). *Speech Analysis for Automatic Speech.*

[12] Nicholas D. Lane, P. G. (n.d.). Can Deep Learning Revolutionize Mobile Sensing.

[13] Nilavra Pathak, M. A. (2015). Acoustic Based Appliance State Identifications for Fine-Grained Energy Analytics. *IEEE International Conference on Pervasive Computing and Communications (PerCom*, 8.

[14] Pavan, G. (2015). pyAudioAnalysis: An Open-Source Python Library for Audio Signal Analysis. *PLoS ONE 10(12): e0144610. doi:10.1371/journal.pone.0144610*.

[15] Sirajum Munir, J. A. (n.d.). FailureSense: Detecting Sensor Failure using Electrical Appliances in the Home. 9.

[16] VARUN CHANDOLA, A. B. (2009). *Anomaly Detection: A Survey.* NewYork: ACM.

[17] Walmart. (n.d.). *walmart shopping store*. Retrieved from www.walmart.com