Universiteit Leiden
Opleiding Computer Science

Diagnosis and Prognosis of Induction Motors and Bearing Using Fast Fourier Transform and Machine Learning Techniques

Master Computer Science, Advanced Computing and Systems

Name: Ewout Zwanenburg

Studentnr: s1542087

Date: 07/05/21

1st supervisor: Prof. dr. T. H. W. Bäck

2nd supervisor: Dr. D. V. Nguyen

MASTER THESIS

Leiden Institute of Advanced Computer Science (LIACS)
Leiden University
Niels Bohrweg 1
2333 CA Leiden
The Netherlands

Abstract

Rotating machinery are essential components in most today's manufacturing and production industries. Therefore, developing efficient approaches for prognostics and health management of rotating machinery has drawn attention of many researchers as well as maintainers. Commonly, approaches used in prognostics for rotating machinery are classified into three types, namely physics-based approaches, data-driven approaches, and hybrid approaches. The physics-based approaches work on the physical knowledge that is acquired by the physical laws. In this approach, a physical/mathematical model for the system or component is developed. Physics-based methods do not require a large amount of data or the data of the failure events. However, to establish this model, a thorough understanding of the physics of the system/component is required. In contrast to physics-based approaches, data-driven methods are much easier to be developed and applied in practical. Data-driven approaches mainly rely on techniques in the field of Artificial Intelligence, which has many ready-to-use tools that could be applied directly with minor modifications. Nonetheless, compared to the physics-based methods, the data driven methods require a large amount of data, including both historical observation and current condition monitoring data. Hybrid prognostics approaches, which are newly developing approaches, aim at integrating the merits of the above-mentioned methods while minimizing their limitations.

In this work/thesis, we propose a method for fault detection and diagnosis of rotating machinery, which combines classical fast Fourier transform and data-driven techniques. We validate our method on two study cases that include induction motors and roller bearings. Our method achieves high accuracy's on real-world data and requires minimal domain knowledge.

Contents

1 Introduction

1.1 Induction Motor and its main defects

Induction motor plays one of the most important roles in the conversion of electrical energy into mechanical energy and it is seen as one of the most fault tolerant and economical motors in the modern industry [1]. This is why many companies tend to use the kind of motors for multiple applications within the industry. The fault tolerance does not mean that there are no defects on the induction motor. The faults that occur can be classified into multiple groups. These groups are the following. The bearing faults, the stator faults, the rotor faults and the other fault that occur within the motor [2]. According to the statistical studies of the IEEE [3], the faults that occurs the most is the bearing fault, this occurs in 42% of the times. The stator fault occurs 28% of the times and the rotor fault occurs 8% of the time. The remaining 22% can be ascribed to the other faults that occur in the induction motor.
The failure of insulation of winding can cause a winding fault, this fault is widely associated with the stator. This fault can lead to heating. This can result in total failure in the stator winding when this fault is not taken care of properly. [4].
The faults concerning the the rotor can be caused by multiple factors. The faults that can occur are for example a broken bar, a broken end-ring, a coupling fault and many more. These faults do not have just a singular cause, they can be caused by many factors. When we look at the difference between medium and small voltage motors the rotor fault can be more dominant due to the thermal stress on the rotors. The broken bar can be caused by multiple reasons. These factor all have to do with and overload of stress. This stress can either be environmental, dynamic, magnetic or thermal [5].
The faults that could be assigned to the fault that were not stated specifically are air-gap eccentricities. These can be dynamic, static or mixed. Between the stator and the rotor could be a non uniform air-gap which could result in the eccentricity errors. These errors can result damage to the core an stator when the friction between these parts become to severe. This could cause the motor to fully stop functioning. The cause of these air-gaps can be ascribed to either a manufacturing error or a bearing fault [5, 6].
The final and most prominent fault is the bearing fault. This, as stated before, covers around 42% of all faults in the electrical motor[3]. This is the reason why this thesis mainly focuses on this fault. The following section will cover this fault in more detail.

1.2 Bearing and its main defects

Bearings are one the most important elements in an induction motor, but it has a lot more applications than just within this induction motor. This bearing is used for rotating different parts within the motor itself. The bearing itself is made out of different parts. The outer ring, the inner ring and a number of rolling elements which are called the balls. These balls can spin within these rings. These most common factors are improper lubrication, the mounting and installation, operational stress and environmental influences [8]. These factor can all result in bearing failure. Even though some of these reasons do not have a clear general cause like the improper mounting but there are some main errors that occur in a functioning induction motor. The bearing errors that turn up in the functioning motors are outer race, inner race, cage and ball errors. These errors can result in multiple forms of dysfunction. The most used methods for failure detection are either the monitoring of the vibration of the motor or the

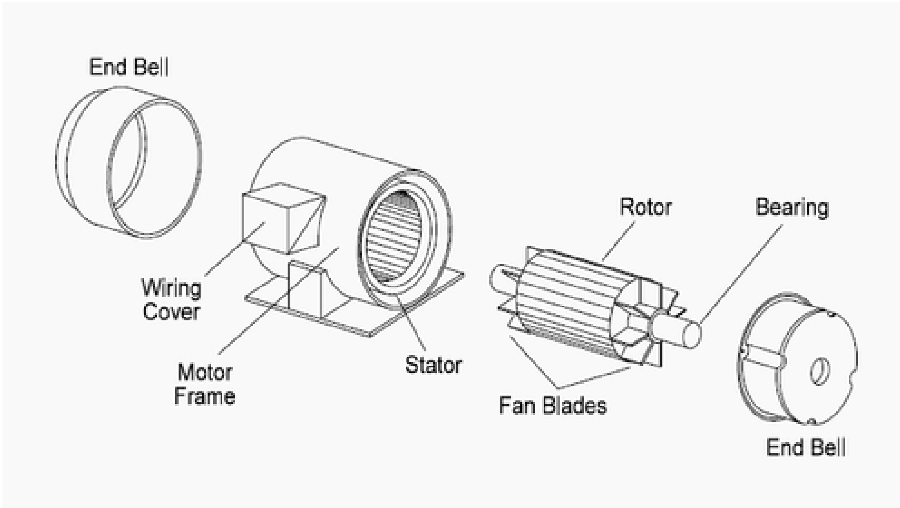


Figure 1: Scheme of induction motor [7]

acoustic emission of the bearing. This is because when a failure occurs it will result in specific vibrations. These vibration can then cause an increase in acoustic emission.

1.3 Review of Diagnosis and Prognosis methods for rotating machinery

Both Prognosis and Diagnosis are measures to calculate the current health of rotating machinery. This can be assessed by calculating this from the symptoms that can be derived from the data that comes from rotating machinery. The three main approaches to this problem are the following. First you can look at fault detection. This is the process of figuring out that a fault has occurred in the motor. Without the actual detection of the fault, no other assessment can be done. When the fault is detected you can either look at fault isolation or fault identification. The isolation means that the fault can be rooted down to a specific part of the motor. The identification is the calculation of what the fault came from. This is important because when you know the root of the fault you could prevent it in the future. The Diagnosis of the rotating machinery means that the rotating machinery will be maintained on the moment when it is most necessary. Unnecessary maintenance is expensive and by looking at past data of broken parts it should be possible to see when an error occurs. This can be done by monitoring the current health of specific parts to see if it needs maintenance at a specific time.
The Prognostics is the calculation of how the element in the rotating machinery will react in the future. This means not only looking at what you can calculate at a specific point but also how it will evolve over time. This makes this a prognostic evaluation of the data. This could result in a prediction called the remaining useful lifetime. This is how long a specific element in the rotating machinery can go on until it breaks down. This together with the classification of the different faults is what this paper mostly focuses on.
There are many methods of either classifying or calculating the remaining useful lifetime. These methods are either physics based, data driven or a hybrid of the two. In A review of physics-based models in prognostics: Application to gears and bearings of rotating machinery [9] different prognostic approaches are discussed considering rotating machinery and bearing. This is all focused on physical models. A data driven approach can be found in different forms. In *An Extension Neural Network and Genetic Algorithm for Bearing Fault Classification* [10]

Neural networks are used to classify the bearing error and its intensity. This paper has found a 100% accuracy with these bearings. It uses traditional Extension Neural Network and improves it by automatic determination of the learning rate. Another method that is used is using multi label classification. This is considered in *Multi-label Classification for Fault Diagnosis of Rotating Electrical Machines* [11]. Multiple methods are considered and compared. This paper uses both vibration and current data. Both these kinds of data will also be considered in this paper.
There is not a main approach to calculating the remaining useful lifetime. This is mainly due to the fact that the variables that the remaining useful lifetime is based upon can either be chosen by the researcher or by an algorithm. In the paper *Predicting remaining useful life of rotating machinery based artificial neural network* [12], neural network is used to calculate the remaining useful lifetime. For these neural network, it uses different features like the time and the fitted measurements Weibull hazard rates of both the kurtosis and the root mean square. This results in a percentage of the expected life total left as the output. In *Remaining Useful Life Prediction for Rotating Machinery Based on Optimal Degradation Indicator* [13] it is stated that the first predicting time and the degradation indicator are subjectively chosen. This is why this thesis is opting to find the optimal degradation factor upon which it uses the Wiener model to calculate the first predicting time. Another approach can be found in models like the bayesian models used in *Remaining useful life estimation of critical components based on Bayesian Approaches* [14]. This shows that neural network is not the only way of calculating the remaining useful lifetime. A method that is widely used within different approaches use the Fast Fourier Transformation, which will be explained in further detail later on. An example of this is *Analysis of the Rolling Element Bearing data set of the Center for Intelligent Maintenance Systems of the University of Cincinnati* [15]. In this thesis the same dataset as the one used in this thesis is analyzed using fast Fourier transformations. This is used to translate the data from the time domain to the frequency domain.
Considering the previous work in comparison with this thesis we can define some limitations to the previous work and how this thesis works with those limitation. Previous work as stated above often use neural networks and other methods which can be seen as expensive operations. These methods are mostly based on purely computer science. This thesis works on a more hybrid approach which aims to combine both the physical and data driven approaches. Another limitations to the previous work is based on fabricated data while this thesis uses real data with both its source in either the current or vibration of the rotating machinery.

1.4 Structure of the thesis

This thesis is structured in the following way. First the general outset of the datasets is given. In this part both datasets will be described in detail. Both the origin of the data and the structure of the data itself will be discussed. Thereafter the outline of the steps taken within the thesis will be discussed in the method section. All different elements in the research will be outlined and explained in detail. After this section the actual results will be discussed and shown. This will be in the same structure as will be explained in the section before. All results will be discussed in detail and the way these results were acquired. The classification and remaining useful lifetime will be discussed in separate sections within this thesis. These sections are section 4.1 and 4.2 Finally both the classification and remaining useful lifetime will be concluded and summarized in separate sections which are 5.1 and 5.2 .

1.5 Research Questions

This thesis will finally conclude in answering the following research questions.

1. How can raw data from the current of a motor be used to classify whether a motor is broken and what error occurred?
2. How can vibration data from bearing be used for the calculation of the remaining useful lifetime?

2 Data Set

This section will describe the two datasets that were used for this thesis. The first dataset is provided by Semiotic Labs Company [16]. This dataset contains data from two induction motors. The second data set is provided by Intelligent Maintenance System. This data contains a test on four bearings that were installed on the same shaft.

2.1 Electrical motors

Data of two induction motors, which are in real working conditions, was provided by Semiotic Labs Company [16]. The data set contains measurements of the motors in 3 different conditions: healthy, coupling fault, and bearing fault.

2.1.1 General Specifications

Two different sets of data were provided by Semoitic Labs Company. These two induction motors have different configurations. The configurations and their technical details are described in Table 1.

2.1.2 Dataset specification

The data consists as stated before of two different datasets. These dataset are split in three different folders. These folders are either healthy, bearing fault or coupling fault. These are the three main sets that belong to these motors. The data consists of real operational data from these motors. The motors start with the error already inflicted on the motors. This is done to make sure that the correct fault occurs in the different test. These adjustments have been made slightly. This means that the motors aren't broken at the very start but it is made sure that over time these faults will occur. The severity of the fault increases over time until the motors either break down or the test is stopped. The motors actually breaking down only occurs with the coupling error. The healthy and bearing error tests are stopped after a finite amount of instances.
The recording duration of each measurement is 15 seconds with a sampling rate of 20 kHz. The instances in the dataset all have a date and two different sets of data. These sets are the current and the voltage at specific points in time. Both current and voltage datasets are split into three columns. These three columns represent the values of the phases that make up the

Id	1	2
Amps	21.2	39.5
Motor type	NK 80-250/270	NK 80-160/167
Power factor	0.86	0.9
Number of poles	4	2
Wiring	wiring.DELTA	wiring.DELTA
Rated voltage	380	380
Efficiency	0.914	0.927
Power	11.0	22.0

Table 1: Dataset motor specifications.

	1	2
Coupling Fault	54	18
Healthy	1477	477
Bearing Fault	2531	10468

Table 2: Amount of files per dataset

three phases system. Both dataset have the same data set up and the same sample rate. The amount of instances for the different datasets are shown in Table 2.
The amount of data that is recorded per instance remains the same for all files. The intervals between instances does vary over time.

2.2 Rolling element bearing dataset

2.2.1 Specification

The rolling element bearing data was created by Intelligent Maintenance System (IMS) and downloaded from [17]. This data contains different experiments using bearings. This data , provided by the center of intelligent Maintenance system, the IMS, contains the vibration data of three different tests on bearings. The following description is taken from the README file included with the datasets. Four bearings were installed on the same shaft. The rotation speed was kept constant at 2000 RPM by an AC motor coupled to the shaft via rub belts. A radial load of 6000 lbs was applied onto the shaft and bearing by a spring mechanism. All bearings were force lubricated. Rexnord ZA-2115 double row bearings were installed on the shaft. PCB 353B33 High Sensitivity Quartz ICP accelerometers were installed on the bearing housing (two accelerometers for each bearing [x- and y-axes] for data set 1, one accelerometer for each bearing for data sets 2 and 3). Sensor placement is also shown in figure 2 . All failures occurred after exceeding designed life time of the bearing which is more than 100 million revolutions.

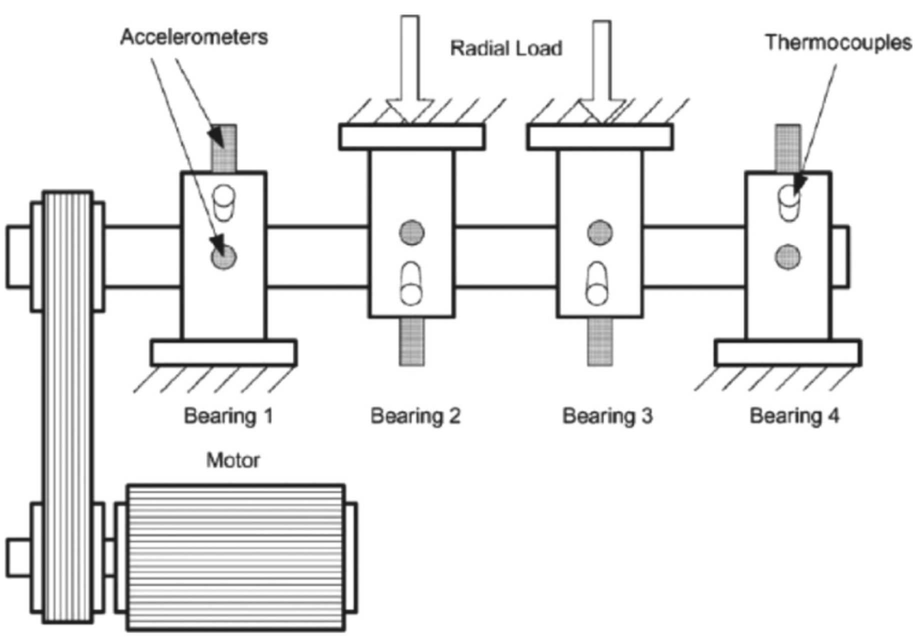


Figure 2: Scheme of IMS test set up [18]

Three (3) data sets are included in the data packet (IMS-Rexnord Bearing Data.zip). Each data set describes a test-to-failure experiment. Each data set consists of individual files that are 1-second vibration signal snapshots recorded at specific intervals. Each file consists of 20,480 points with the sampling rate set at 20 kHz. The file name indicates when the data was collected. Each record (row) in the data file is a data point. Data collection was facilitated by NI DAQ Card 6062E. Larger intervals of time stamps (showed in file names) indicate resumption of the experiment in the next working day.[17]

First test	
Recording Duration	October 22, 2003 12:06:24 to November 25, 2003 23:39:56
No. of Files	2,156
No. of Channels	8
Channel Arragement	Bearing 1 – Ch 1 and 2; Bearing 2 – Ch 3 and 4; Bearing 3 – Ch 5 and 6; Bearing 4 – Ch 7 and 8.
File Recording Interval	Every 10 minutes (except the first 43 files were taken every 5 minutes)
File Format	ASCII
Description	At the end of the test-to-failure experiment, inner race defect occurred in bearing 3 and roller element defect in bearing 4.

Table 3: Description First Test

Second test	
Recording Duration	February 12, 2004 10:32:39 to February 19, 2004 06:22:39
No. of Files	984
No. of Channels	4
Channel Arragement	Bearing 1 – Ch 1; Bearing 2 – Ch 2; Bearing 3 – Ch 3; Bearing 4 – Ch 4.
File Recording Interval	Every 10 minutes
File Format	ASCII
Description	At the end of the test-to-failure experiment, outer race failure occurred in bearing 1.

Table 4: Description Second Test

Third test	
Recording Duration	March 4, 2004 09:27:46 to April 4, 2004 19:01:57
No. of Files	4,448
No. of Channels	4
Channel Arragement	Bearing 1 – Ch 1; Bearing 2 – Ch 2; Bearing 3 – Ch 3; Bearing 4 – Ch 4.
File Recording Interval	Every 10 minutes
File Format	ASCII
Description	At the end of the test-to-failure experiment, outer race failure occurred in bearing 3.

Table 5: Description Third Test

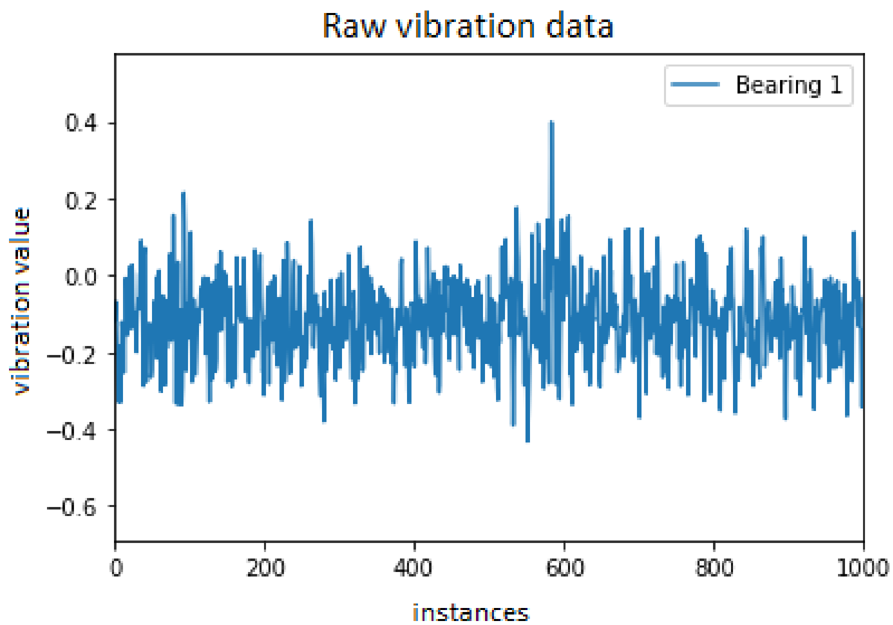


Figure 3: Part of the Vibration signal of the IMS data

2.2.2 Vibration

When considering the vibration of the bearings there is a special device which registers the different amplitudes of the vibration in one specific direction. For one bearing the registration was done in two directions but because this hasn't been done for all bearings we only consider one direction which is used for all the other bearings as well. By using this amplitude we can see what vibrations were in the bearings at those specific windows. Not all measurements were done directly after each other. The motor was stopped at some points in time and reinstalled again. In this thesis we don't look at these open spots but consider the amount of cycles we read rather than the exact moment the motor was on or off. The following figure 3 shows a small sample of the vibration data.

2.2.3 Error Harmonics

In the following table the frequencies of the different errors are stated. These values are found in [15].
The errors we need to consider for this thesis are the Ball Pass Frequency Outer Race and the Ball Pass Frequency Inner Race. The first data set has an inner race error in bearing three and

Characteristic frequencies	
Shaft frequency	33.3 Hz
Ball Pass Frequency Outer Race	236 Hz
Ball Pass Frequency Inner Race	297 Hz
Ball Spin Frequency	278 Hz (2 x 139 Hz)
Fundamental Train Frequency	15 Hz

Table 6: Characteristic frequencies [15]

the other two sets have an outer race error in bearing one for set two and in bearing three for set three. The error harmonic means that instead of only looking at the given frequencies we also look at the harmonic of these frequencies. This means that we look at multiplications of these frequencies and also use those for the energy that we calculate.

3 Method

3.1 Data exploration

Data exploration is the process of exploring the dataset that is given to the scientist. This can be done in different ways. This is the important first step in the analysis of the dataset that is used in research. The data exploration is done before any analysis is done to give the researcher more insight in what the dataset contains. This is done to prevent the error of any forethought knowledgde of the dataset that turns out to be false. The data exploration is mostly focused on visualizing the given dataset to understand it better [19]. The visualization of the dataset is an important step in understanding the dataset used by a researcher. Table and number are abstract to us. The visualization used for the data exploration is important for understanding the data, rather then examening the number by themselves. The chosen visualization is therefore considerable to show the right patterns that are within the data. This has also been done in this thesis. A visualization is used for understanding the raw data and examining specific patterns in the data.

3.2 Data Processing

Data processing is the transformation of the dataset used by the researcher. The most important reasons of data processing can be summarized as boosting the consistency of the data and smoothing the dataset so the analysis that can be done easier.This process can include multiple steps. These steps can be summarized in the following main steps: data cleaning, data transformation and data reduction [20]. Data cleaning is the first step that can be taken in the data processing. This step is mostly considered when you are uncertain of the reliability of the given dataset. By looking at different factors like missing or noisy data, processing the data before working on the analysis can be advantages. Data transformation can be done in numerous ways and forms. The main goal of data transformation is to translate the data to a form that can be used for better analysis. Normalization and generalization are of many different approaches to transforming the data[20]. This research will use multiple data transformation approaches to the raw dataset. This includes fast Fourier transformations and pre whitening methods which will be explained later on in this section. Data reduction also consists of numerous different methods of reducing the data for analysis. The reasoning behind this is mostly due to the amount of data that will be analyzed. A big dataset can be very costly to analyse in total while an analysis of the reduced dataset can give insights on the whole dataset. Methods like feature selection can be used to reduce the amount of data that is analyzed [20].

3.2.1 Fast Fourier Transformation

Fast Fourier transformation is a fast low complexity algorithm that calculates the discrete Fourier transformation of a specific dataset [21]. This transformation is one of the most important algorithms in modern mathematical and computer science because of its applicability to one and multidimensional systems theory and signal processing [21].
Fourier transformation is the operation of transforming data in the time domain to the frequency domain. This can be used to analyze specific signals and their frequencies. These frequencies will all be decomposed for the time series into distinct frequencies which are present in the data. The discrete Fourier transformation formula can be written as shown in equation 1.

$$x[k] = \sum_{n=0}^{N-1} x[n]e\frac{-2j\pi kn}{N}$$

Equation 1

Here the $x[n]$ is the discrete signal where $x[k]$ is the resulted Fourier transform. N is the domain size. The Discrete Fourier transform is calculated by multiplying all its values by e raised to a function of n. The result is then summed up for a specific n. The complexity of this algorithm is $O(N^2)$ [22].
The formula can be separated in both an even and an odd sub-sequence. When this is done both the odd and even sub-sequences can be computed at the same time because of equation 2.

$$x[k] = x_{even}[k] + e\frac{-2j\pi kn}{N}x_{odd}[k]$$

Equation 2

The equation given above, because of its separation of the even and odd sub-sequences, can be calculated faster with the complexity of $O(Nlog_2N)$ [22].

These formulas will be used on the dataset to calculate what frequencies are present at specific points in time and in what amplitude.

3.2.2 Feature extraction

Feature extraction is the art of using the features from your dataset and translate them to new features. The features that the feature extraction works on are either the features of the raw data of features that were calculated before. With a big dataset the amount of features could be enormous. So calculating more insightful features from this big stack of features could be beneficial to the analysis. The most important benefits of the feature extraction are accuracy improvements, over fitting risk reduction, speed up in training, improved data visualization and increase in explainability of the model that is created [23]. This research will mainly use feature extraction in the form of accuracy improvement and improved data visualization.

3.2.3 Pre-whitening

A spectrogram is a form of visualization on which it is possible to see the change of specific features over time. These features have a specific value and this value can increase or decrease over time. In the spectrogram this can be visualized by coloring these values. The change of color can show how the features change. With pre-whitening the overall spectrogram can be normalized by its first instances. This is can result in clearer images of the differences over time. This is due to the fact that specific features could have consistently high values that could overshadow lower values. By normalizing the data by the first instances the change in values can be spotted more clearly over time because high constants are filtered out. The contrast is therefore enhanced [15]. This method will be used for the visualization of the frequency changes.

3.2.4 Data imbalance

Data imbalance means that the amount of instances in two classification have a significant difference between one another. This would mean that we have loads of instances for classification A but very few for classification B. Data imbalance can be dealt with in multiple ways. The methods used in this thesis are resampling and the Synthetic minority oversampling technique. Resampling is the method of copying the the existing data that is used and multiplying it so that the amount of data is changed to the size of the larger dataset. This method does not create new instances but rather copies them from existing ones [24]. Synthetic minority oversampling is the method that takes the minority dataset and creates new instances that are similar to the existing instances. It uses a nearest neighbour approach to calculate these new instances. By this method a new point in the decision space is chosen. Now a line can be drawn from the original to the newly chosen spot by the k nearest neighbour algorithm. A convex combination is used to generate the new instances used for the imbalance. The number of created instances can be made equal to the amount of instances in the majority dataset [24]. For this part of the method we will look into the different algorithms which are used within this thesis for classification of the different stages of the motor. The algorithms which are used are some of the most widely used algorithms for data classification. I will discuss them now briefly.

3.2.5 Algorithms

For the machine learning algorithms the sci-kit learn package was used. This package contains all basic machine learning algorithms that are mostly used for data analysis. These algorithms are easy to use using the functions that are given within the sklearn package, such as random forest and support vector machines. There is also a lot of extra information online on how the package works which makes this the best option to use for your machine learning experiments. [25]

3.2.5.1 Decision Tree

The decision tree is the most straight forward algorithm used for this research. The algorithm learns from the given features what he can split the data from. So for example when a certain peak is higher than a specific value it belongs to one group and when it's lower it belongs to the other group. This way the algorithm keeps splitting the data until only specific groups are left which belong to either the healthy group, coupling error or bearing error. These decisions are based on the given training set and it will fit the tree as good as possible on the given data. This results in the decision tree of that data set. When we try to optimize the results of the decision tree we look at multiple hyper parameters for this algorithm. These can all be chosen using a grid search. This grid search looks at multiple sets of combination of different hyper parameters and calculates the result of the algorithm using those specific hyper parameters. This way a vast amount of different set ups is considered to find the optimal hyper parameters for this algorithm. The grid search takes as much hyper parameters into account as possible. Some of which are the depth of the tree, the amount of features used and the splitter, which could be best or random and there are many more hyper parameters to be tuned. [26]

3.2.5.2 Random Forest

The random forest algorithm is mostly based on the algorithm stated above. This algorithm makes a finite amount of decision trees and classifies the instance according to their outcome. The way this algorithm does that is based on the outcome that comes up the most by the different algorithms. Because these trees all grew on their own they all have their own strengths and weaknesses. The reason the random forest algorithm works is because all these trees grow independently the individual errors of these trees should be neglected because they all won't make the same error. So the forest has a better overall outcome than the individual trees. Random forest has a lot of hyper parameters which are equal to the decision tree like the maximum depth and the amount of features that are considered. The most important feature that is new to the random forest is the amount of trees that are created by the algorithm. It is important to find a good amount of trees for the algorithm to function properly. This is why there are many different amounts of trees considered in the grid search. [27]

3.2.5.3 Logistic Regression

Logistic regression is the way that the algorithm calculates what the probability would be that an instance would be in one class or the other. Instead of saying 0 for not in the class and 1 for in the class, the probability could be 0.8 which would indicate that the chance that is is in the class is quite high. After these probabilities are calculated a threshold could be calculated or a specific line using linear regression. This can then split the calculated values in the class or not. The hyper parameters used in this method aren't as impactful as for the other methods but there are still some parameters to be considered. The most important ones are the solver, which can be done by a number of different algorithms, the penalty given to errors and the C parameter which determines the severity of the penalty calculated by the penalty algorithm. Even though these parameters are less impactful they could still prove to show improvements. [28]

3.2.5.4 Support Vector Machine

A support vector machine is based upon the distinction between the classes in a hyperplane. A hyperplane is a grid with multiple dimensions. For humans two or three dimensions is easy to visualize but beyond three it gets difficult. The support vector machine tries to distinct the different classes as far as possible from each other in specific hyper planes. It tries to create a boundary which is as far as possible from the point that is the closest within both sets. The higher the distance between the boundary and the closest point, the higher the confidence to say that a new point that falls within these planes is a specific class. Some of the most important hyper parameters to be tuned for the support vector machines are the range between the classifying line and the classified point and the penalty given for wrongly classified points. This will also be done using the grid search method discussed earlier [29].

3.3 Health Index

The health index is a factor that states the condition of a specific subject. This subject could be anything like a machine to a human being. The health index shows what the current condition of the subject is. This condition is derived by multiple factors. These factors vary for different subjects. These factors have to be chosen from the features acquired from the given dataset.

The features are chosen by the researcher. Every subject could have different features which affect the health index. When the features are combined this should result in a single index that decreases over time. When the health index is followed over the whole life cycle of the subject the line should eventually hit zero. This is the point where the subject fails. In the case of rotating machinery this is the moment the machine stops working. For convenience, researchers/users normally set the health index between 1 and 0; 1 and 0 indicates the totally healthy subject and totally faulty subject, respectively.

3.4 Remaining Useful Lifetime

The remaining useful lifetime for a specific subject is the time that is left until the subject eventually fails. The subject in this case can also, just like the health index, be anything that has a limited life cycle. The remaining useful lifetime uses the health index that was explained in the previous section and predicts how this index will decrease over time. When only the health index in considered we would only know the current status. By calculating the remaining useful lifetime we can predict how this status will move on in the future. This can be done in multiple ways of prediction. This can vary from simply continuing the current health index as it goes until a full time series analysis. The goal of remaining useful lifetime is to calculated when the subject will fail so you can be prepared when it eventually fails or maintain the subject before it will fail.

4 Results

4.1 Classification

This section will go into the process of classifying the dataset received from Semiotic Labs into healthy, coupling error and bearing error. The process will of classification will be outlined from acquiring the data to the final results.

4.1.1 Data exploration

The data exploration is done by creating a visualization of the raw data that was received. This data was, as told before in section 3.1.2, split into three phases that make up the three phase system. In the following figures an example visualization is shown of the dataset. The three different lines represent the three different phases. This is both shown for the current and voltage, see figure 6 and figure 7.
It is clearly visible that these lines show a clear image of the three phase system of both the current and the voltage. All three lines show the same pattern over the time that the test measurement is running. The three lines also all show the same anomalies that are clearly visible in the form of peaks and fall. These anomalies are all very consistent. This is probably due to the fact that the energy source is a variable supply unit. The test set up it not directly connected to the electrical grid. This is done so that the supply frequency can be changed during the test.

4.1.2 Fast Fourier Transformation

We converted the induced current from time domain to frequency domain using FFT. The current was chosen as the main dataset because the current is more expressive compared to the bare voltage. This is because the current can be seen as the function of the voltage.

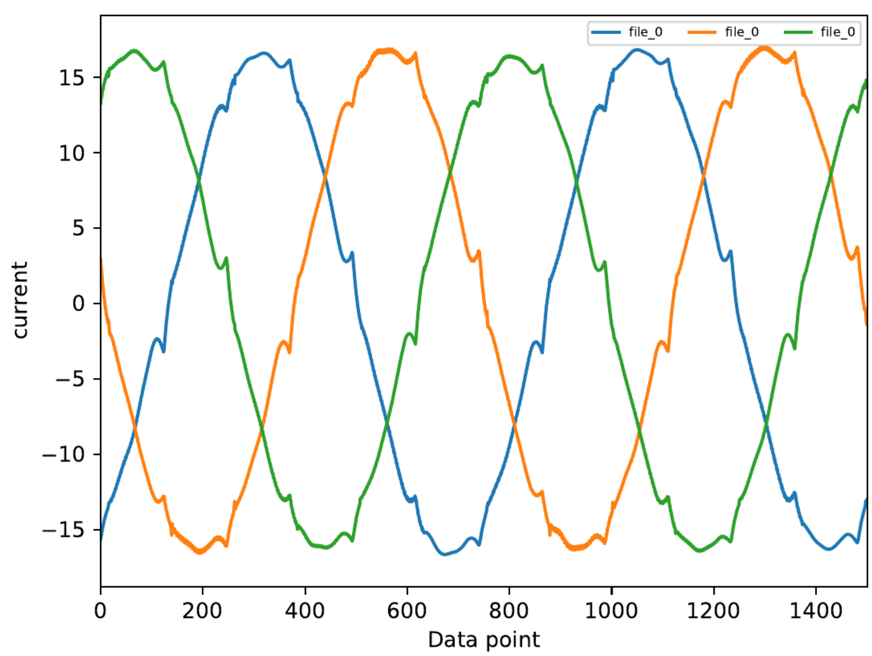


Figure 6: Part of the current signal of motor 1

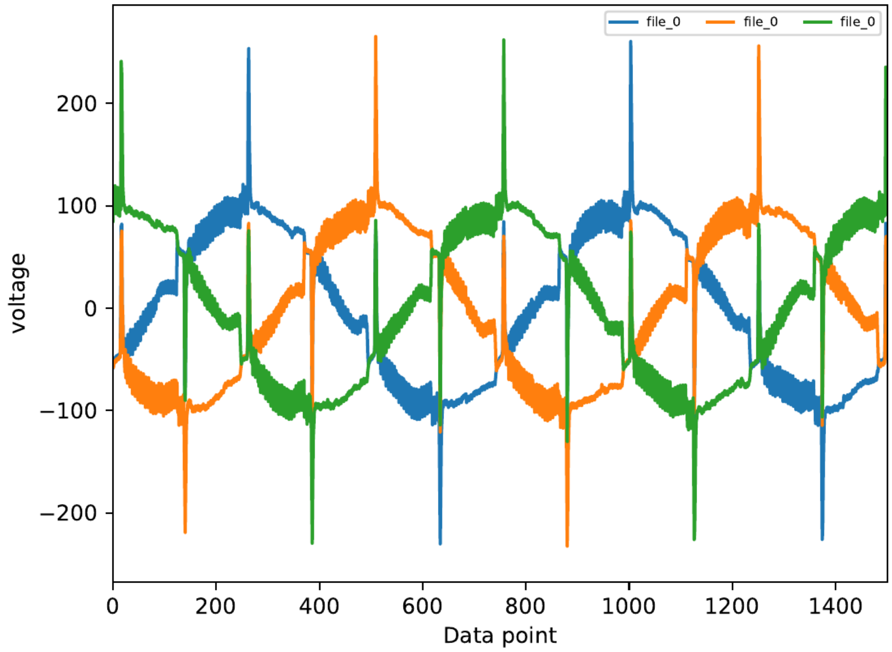


Figure 7: Part of the voltage signal of 1

The current data consists of three phases. These phases were averaged before calculating the fast Fourier transformation, short cut FFT. These different instances were then all translated to FFTs. These FFTs show as stated before what frequencies were present in the different instances. The FFTs were done using the scipy package [30]. All calculated FFTs were then combined into one big data set which could then be used for further research. This was done for all three different datasets of both motors. figure 8 shows an example of the average FFT spectrum of different motor conditions including healthy, bearing fault, and coupling fault.

The fast Fourier transformation was done on the raw data. The current was chosen as the main dataset because the current is more expressive compared to the bare voltage. This is because the current can be seen as the function of the voltage. All different files were transformed into fast Fourier transformations in the same manner using the scipy package [30]. This package can be used on time series datasets and produce the corresponding fast fFourier transformation. This fast Fourier transformation is then calculated to a single column with the amplitude for each individual point in the frequency domain. This shows whether the frequency is present in the time series data or not. All these columns were then combined into one large dataset with columns corresponding to the files they came from. This is done for the healthy, coupling and bearing datasets. figure 8 shows an example of fast Fourier transformation for all three different dataset types combined in one graph.
In figure 8 we can see that a clear main frequency of 50 Hz is present in the data which also results in peaks around 150 Hz and 250 Hz. The peak before 50 Hz can be mostly written down to noise in the data as all three dataset showed similar abnormalities around 25 Hz. This graph also shows that the three dataset show a similar fast Fourier transformation apart from some difference in amplitude. The difference in amplitude of 150 Hz does not immediately show the bearing fault because the amplitudes in all three datasets vary between different instances.
From this visualization can be concluded that the main peak of 50 Hz and its harmonics are

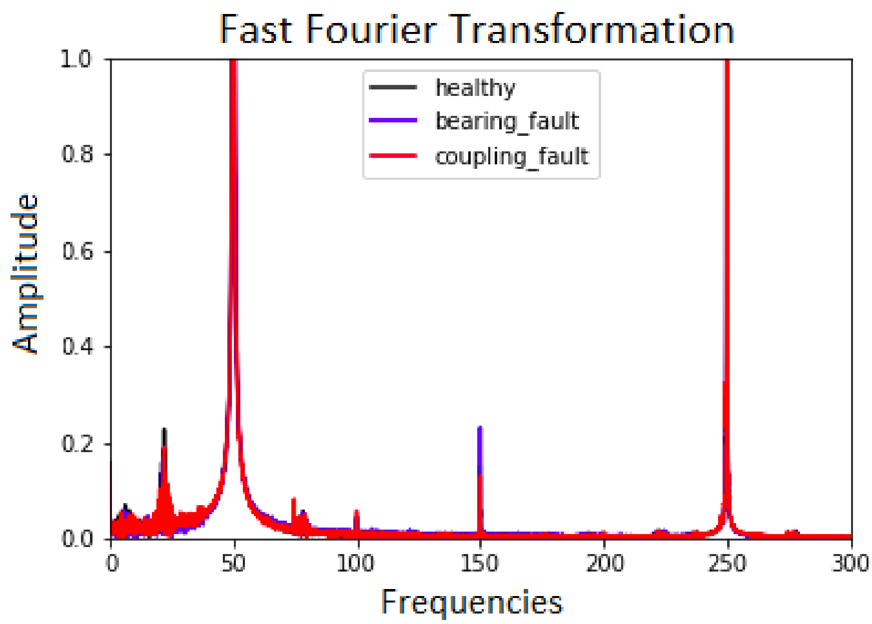


Figure 8: Fast Fourier Transformation showing all three conditions

the biggest peaks present in the raw dataset. The peak that can be seen before the main peak of 50 Hz can be assigned to misalignment in the set up and noice in the dataset itself.

4.1.3 Feature extraction

Feature extraction was done on the fast Fourier transformations calculated in the previous section. The main frequency was chosen as the base for this thesis. The dataset that was analyzed mostly consisted of data that was acquired with the main frequency of 50 Hz. As told before the energy supply was variable, this shows in the data in the way that some instances have a main frequency of 25 Hz. These fast Fourier spectra with different main frequencies are too different to analyze in the same manner. This resulted in the fact that we only used data with a main frequency of 50 Hz because most data had this main frequency. The data with the main frequency of 25 Hz was not considered and removed from the dataset.
The energy is a feature that is calculated by summing up the amplitudes within a specific window. By looking at this feature over time it can be shown how the amplitude, and thus the presence of a specific frequency, changes over time. The windows that were chosen for this feature are around the main peaks and between them. This can be written down as the windows $[0-40][40-60][60-90][90-110][...]$. This resulted in 80 features.This can be visually shown in figure 9.
Here the window taken around the main peak of 50 Hz is shown. Within these windows the amplitudes were summed up for every distinct window and turned into a new dataset that is used as the feature set.
The method shows a rather simple approach to the feature extraction issue. Rather than using different forms of features, which is done in previous papers as well, this single feature extraction methods proves to be very expressive over the different windows in the dataset.

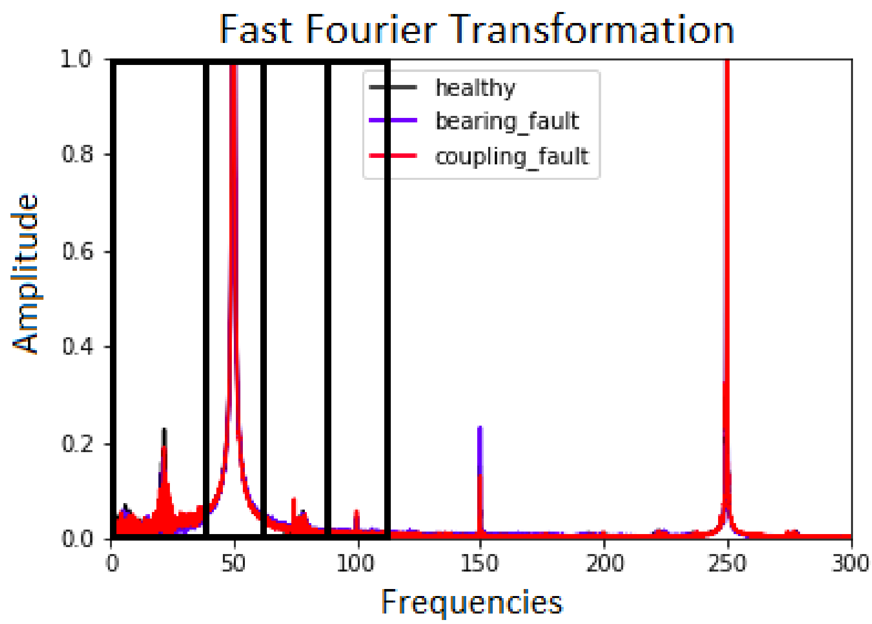


Figure 9: Chosen energy window for feature extraction.

4.1.4 Machine learning results

This section will go into the classification of whether the data is healthy or unhealthy considering both coupling and bearing faults. After this classification the unhealthy data will be classified as either the coupling fault or the bearing fault.

4.1.4.1 Healthy versus unhealthy classification

For the classification of the healthy versus unhealthy data both the coupling fault and bearing fault datasets are combined in the unhealthy dataset set. The features, as stated before, are the energy calculated from the windows chosen that were described in section 4.1.3. A gridsearch is used to optimize the hyperparameters for the different algorithms. The machine learning models used for this classification are the decision tree, random forest, logistic regression and support vector machines. These classification models were used because these models are some of the most well known models used for research. These have all been tested numorous times and often show good results even though they are quite simple to use and understand. The first results that were acquired with this method are shown in Table 7
We see here that the accuracy is of the models are overall very good but when we look at the

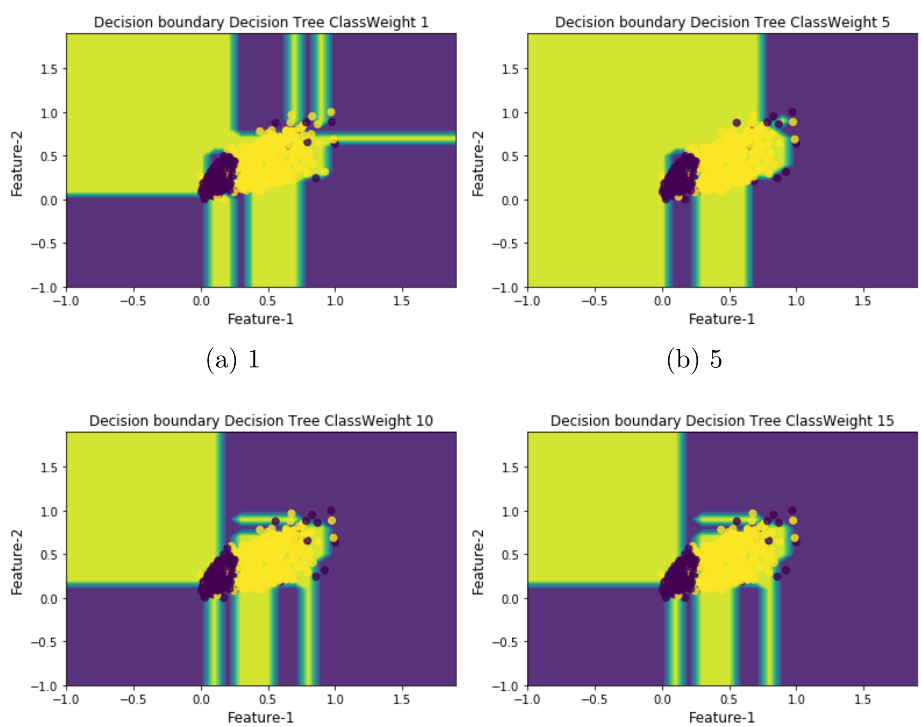
Algorithm	Class Weight	Accuracy	Precision	Recall	F1	False Neg.
Logistic Regression	1	0.9809	0.98	0.99	0.99	17
Support Vector Machine	1	0.9611	0.96	0.98	0.97	42
Random Forest	1	0.9776	0.99	0.98	0.98	15
Decision Tree	1	0.9683	0.98	0.97	0.98	17

Table 7: Results for different methods, class weight 1

amount of False Negatives they are still quite high. False negatives mean that an error is not classified as an error. This can be very costly when an error is not found. This is why a lower amount of false negatives is desired. This could result in an increase in True negatives. These are healthy conditions which are classified as unhealthy. This means that the cost of maintanence in increased because there are more false flags of errors. This is a tradeoff which can be made by the user themselves. The models can be adjusted to reduce the amount of false negatives using the class weight.The class weight is used to put more empathize on a specific class [31]. This is used to more empathize on the unhealthy class. This is done because this makes sure that an instance is classified as unhealthy more frequently because of its higher weight in the models. The results of the four different algorithms and their decision boundaries are shown in figures 10, 11, 12 and 13. These results are received by only considering the dataset of motor 1.

Figures 10, 11, 12 and 13 show that by increasing the class weight the areas that are classified as unhealthy are more narrowed. This shows for two out of four of these algorithms. The random Forest shows the opposite result. It shows a decrease in False Positives by increasing the area that is labelled as unhealthy. This is the same for the decision tree.

The features that were used can be ranked in the order of their importance. The importance of features is a value that shows how important specific features were to the final classification. The following 10 features were the most important from the feature set. These values are based on the Random Forest algorithm.The top 10 of the most important features, based on the Random Forest algorithm, are shown in figure 14.

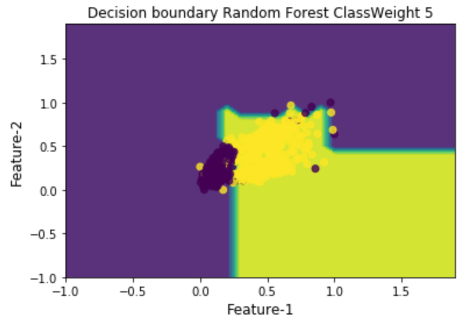


(a) 1 (b) 5

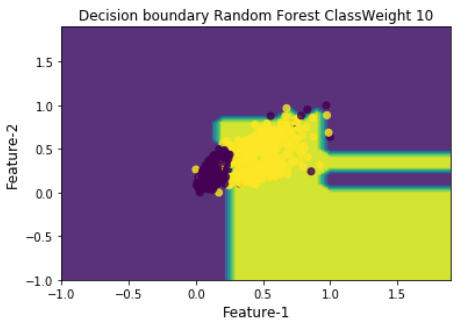
(c) 10 (d) 15

Figure 10: Decision boundary of Decision Tree classifier with classweight

(a) 1



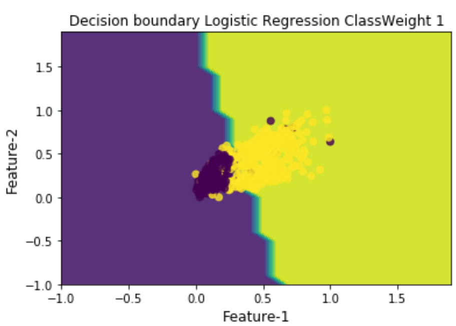
(b) 5



(c) 10

(d) 15

Figure 11: Decision boundary of Random Forest classifier with classweight



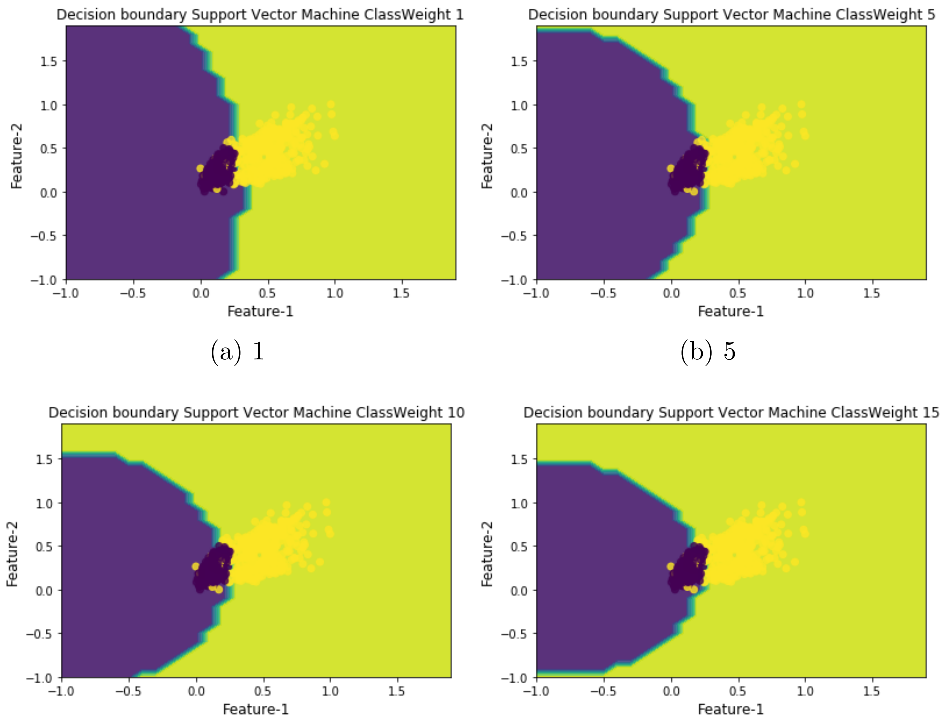
(a) 1

(b) 5

(c) 10

(d) 15

Figure 12: Decision boundary of Logistic Regression classifier with classweight



(a) 1

(b) 5

(c) 10

(d) 15

Figure 13: Decision boundary of Support Vector Machine classifier with classweight

The results shown earlier are calculated by only considering the data of motor 1. To show that this method can be used on different motors figure 15 shows the results of the increase of class weight for the combined data of motor 1 and 2. The solid line shows the amount of False

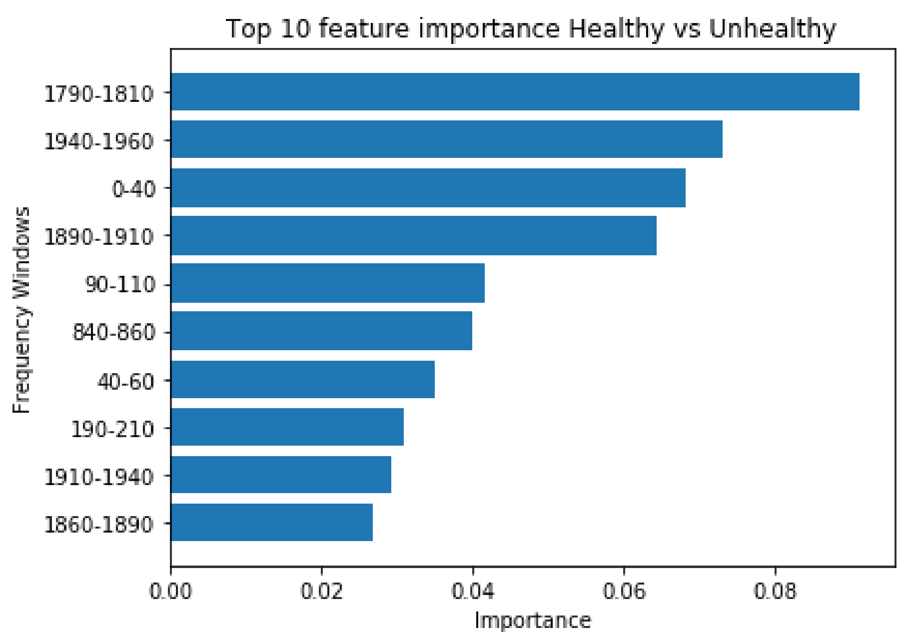


Figure 14: Feature importance healthy versus unhealthy

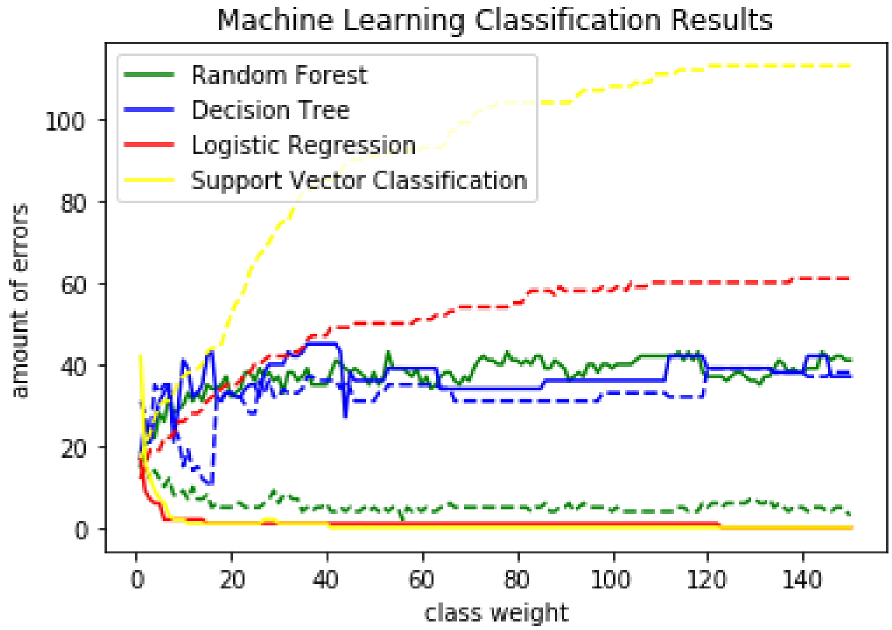


Figure 15: Result of different Classification algorithms

Negatives. The dotted line shows the amount of false positives. A False negative in this graph means that a fault is not found. The false positive means that a healthy machine is flagged as an unhealthy machine. We want to reduce the amount of False Negatives because these errors in the classification can do the most harm to the motor. When an error is not found while it is present, the motor could break down even further. This is why the amount of False Negatives needs to be reduced to a minimum.
Figure 15 shows that by increasing the class weight both the logistic regression and the support vector machine have almost no False Negatives left. The amount of false positives does increase. In the Table below the exact results are shown for the different algorithms.
From these results can be concluded that using the energy as a main feature for the classification does show good results. This method, combining physical and data driven knowledge, does not use heavy neural networks to classify the faults but rather basic machine learning methods. This proves that heavy calculations are not required for a good classification, which has mostly been done in previous work. By tuning the class weight the amount of False Negatives can even be reduced to a minimum. This can be seen when we look at the results of both Logistic Regression and the Support Vector Machine. The amount of False Negatives were reduced to almost zero.
When considering the different classification methods the Logistic regression and Support vector machine outperform the Decision Tree and Random Forest. This is probably because of

Algorithm	Class Weight	Accuracy	Precision	Recall	F1	False Neg.
Logistic Regression	15	0.9789	1.00	0.97	0.99	1
Support Vector Machine	15	0.9710	1.00	0.96	0.98	1
Random Forest	1	0.9776	0.99	0.98	0.98	15
Decision Tree	1	0.9683	0.98	0.97	0.98	17

Table 8: Best results for different methods.

the nature of the problem. The classification is a two classes classification. With this kind of classifications Support Vector Machines and Logistic Regression normally work better because these algorithms try to split the decision space as good as possible while the other methods try to make decisions on specific values of the features. Because of the amount of False Positives we can say that from the two best performing algorithms Logistic Regression performs the best, which is also shown in Table 8.

4.1.4.2 Bearing fault versus Coupling fault classification

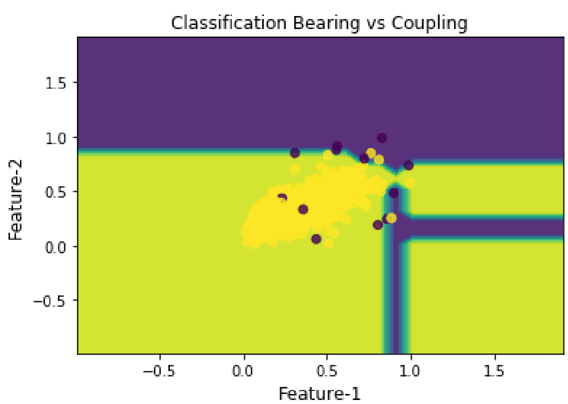
The bearing and coupling feature sets are as stated before imbalanced. The bearing fault dataset is far larger than the coupling fault dataset. These first results show the classification of the coupling fault versus the bearing fault on the dataset of motor 1. The results are received by checking the model using a test set created from the feature set. This is why in the result there are very few coupling instances while the dataset is scaled up later on using the SMOTE and resampling methods.
These tables all show high accuracy considering the classification of the motor 1 failures. To prove that this method is not overfitted on this specific motor the following results are from the datasets of 1 and 2 combined.
Figures 16 give the same classification but also show how the decision lines are created using the random forest algorithm for this classification.
For the classification of the bearing fault versus the coupling fault we can also show what features were the most important to the final classification. We do this in the same way as we did with the healthy versus unhealthy classification by ordering them by their importance. These values are also from the Random Forest algorithm. The result can be seen in figure 17
By looking at these results we can state that it is possible to classify the fault when a fault is found in the data. With the combined datasets it is shown that the errors that occur are mostly predicted as bearing errors due to the large data imbalance. This imbalance can be

Algoritm	Imbalance method	AUC	Precision	Recall	F1-Score
Decision Tree	Original set	0.999	0.93	1,00	0.97
	Resampling	0.963	0.93	0.93	0.93
	Oversampling	0.963	0.93	0.93	0.93
	Over and Undersampling	0.999	0.93	1.00	0.97
Random Forest	Original set	1.00	0.93	1.00	0.97
	Resampling	1.00	0.93	1.00	0.97
	Oversampling	1.00	0.93	1.00	0.97
	Over and Undersampling	1.00	0.93	1.00	0.97
Logistic Regression	Original set	0.998	0.93	1.00	0.97
	Resampling	0.998	0.93	1.00	0.97
	Oversampling	0.998	0.93	1.00	0.97
	Over and Undersampling	0.999	0.93	1.00	0.97
Support Vector Machine	Original set	-	0.93	1.00	0.97
	Resampling	-	0.93	1.00	0.97
	Oversampling	-	0.93	1.00	0.97
	Over and Undersampling	-	0.93	1.00	0.97

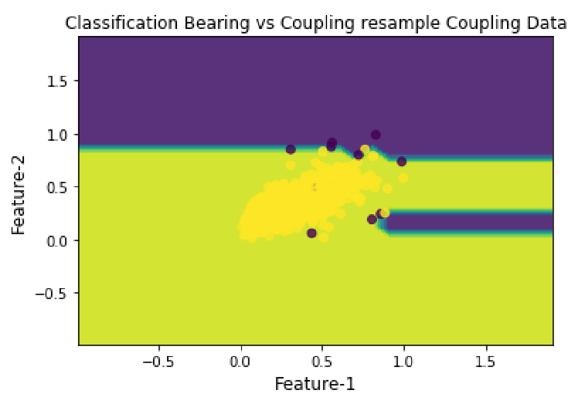
Table 9: Classification results different imbalance methods, including Resampling, Over(sampling) and Under(Sampling) motor 1

Algoritm	Imbalance method	AUC	Precision	Recall	F1-Score
Decision Tree	Original set	0.858	0.68	0.72	0.70
	Resampling	0.859	0.76	0.72	0.74
	Oversampling	0.885	0.64	0.78	0.70
	Over and Undersampling	0.878	0.38	0.78	0.51
Random Forest	Original set	0.999	1.00	0.72	0.84
	Resampling	0.999	1.00	0.61	0.76
	Oversampling	1.00	0.94	0.89	0.91
	Over and Undersampling	1.00	0.94	0.94	0.94
Logistic Regression	Original set	1.00	1.00	0.67	0.80
	Resampling	1.00	0.90	1.00	0.95
	Oversampling	1.00	0.95	1.00	0.97
	Over and Undersampling	1.00	0.94	0.94	0.94
Support Vector Machine	Original set	-	0.90	1.00	0.95
	Resampling	-	1.00	1.00	1.00
	Oversampling	-	1.00	1.00	1.00
	Over and Undersampling	-	0.95	1.00	0.97

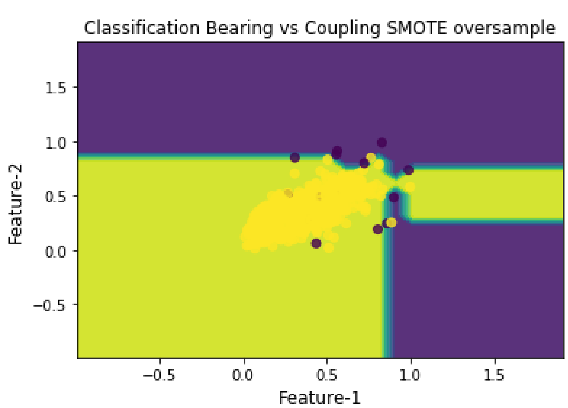
Table 10: Classification results different imbalance methods, including Resampling, Over(sampling) and Under(Sampling) motor 1 and 2



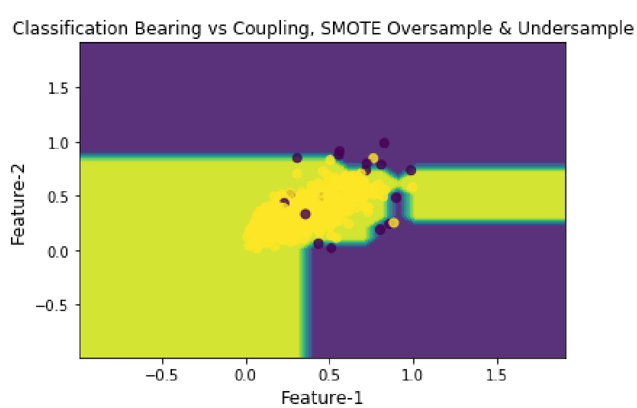
(a) original



(b) oversample



(c) SMOTE oversampling



(d) SMOTE oversampling and undersampling

Figure 16: Decision boundary of Random Forest methods using data imbalance method

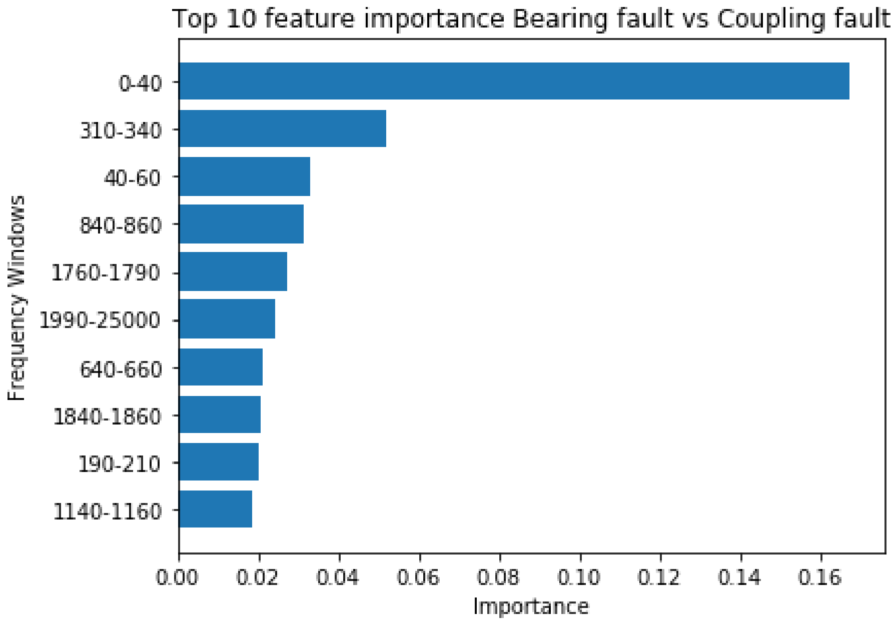


Figure 17: Feature importance bearing fault versus coupling fault

improved by using SMOTE. Both show slight improvements over the overall dataset. This is not the case with the 1 alone. The differences are small but when considering the decision boundaries in figure 16 we see that the more decision space is considered as coupling the better the classification result is. This could prove that bearing fault have a more consistent output while coupling has multiple outputs in different areas in the decision matrix. It has to be stated that the results given by both SMOTE methods vary over different random seeds. This is because the instances that are created are random as stated in the method section. This makes it so that both the original and normal oversampling results remain the same while the SMOTE results vary.

4.2 Remaining Useful Lifetime prediction of machine bearing

In the following section the remaining useful lifetime will be calculated using the Center of Intelligent Maintenance Systems dataset. This dataset shows a full lifecycle of different bearings which can be perfectly used to calculate a health index and eventually a remaining useful lifetime. First the dataset will be explored to see how the data is built up. Then the remaining useful lifetime will be calculated in different steps.

4.2.1 Fast Fourier Transformation

The Center of Intelligent Maintenance Systems dataset consists of different datasets of multiple bearings. There are three main datasets that consist of data from four bearings. This data is the measure of vibration of these different bearings. The Fast Fourier Transformation was also used to translate the data to the frequency spectrum for this dataset. The FFT can then be used to do the analysis on the dataset. These FFT's were, same as for the previous dataset, created with the scipy package [30].

4.2.2 Energy Increase

For the analysis of the different FFT's features need to be created from these FFT's to actually do the analysis on. There are multiple features that can be derived from the FFT's. In the previous section of this thesis the energy was the main feature to classify the dataset. This feature, as stated earlier, calculates the combined amplitude of specific windows of the FFT's. The classification showed very promising results. This shows that energy is a very helpful feature. This is why in this section the energy will also be used as the main feature to show that the energy might also be a good feature for the calculation of the remaining useful lifetime the following example will be explored in figure 18. This example shows an example of how the energy behaves over time.

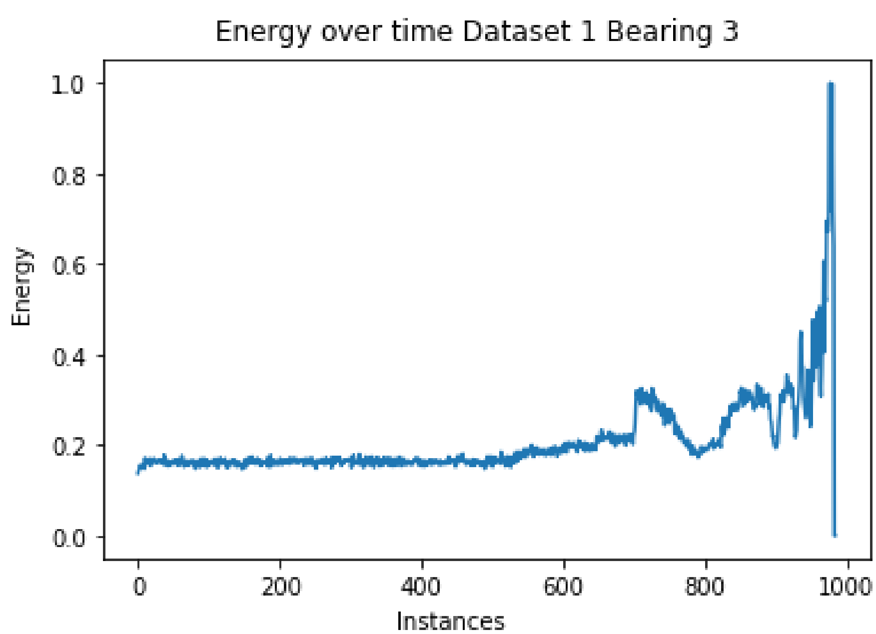


Figure 18: Example of an increasing energy window

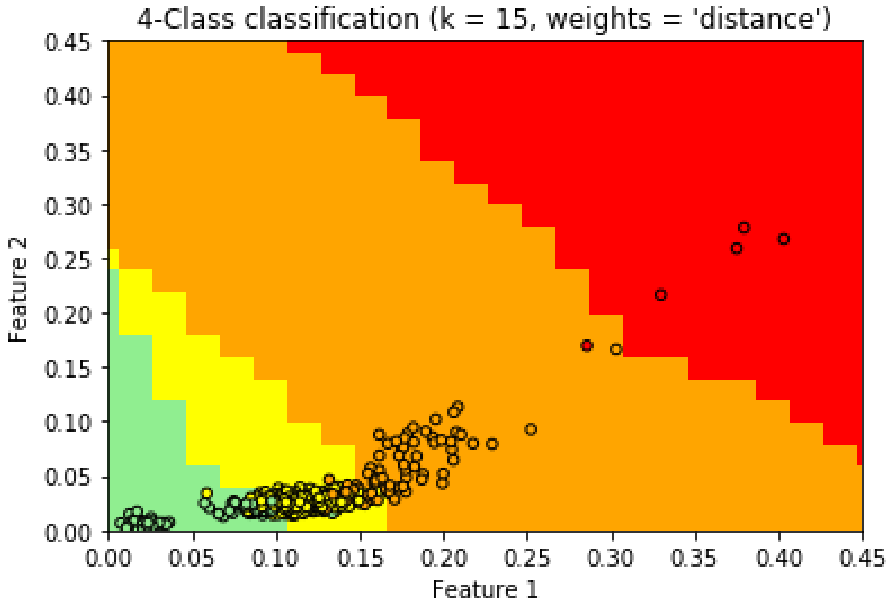


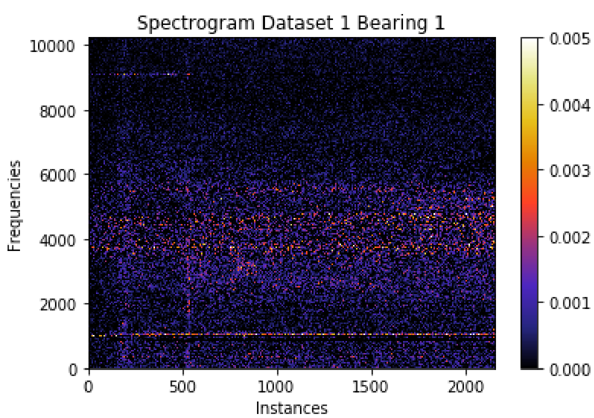
Figure 19: Scatterplot of the different phases of the bearing 3 of dataset 1

The example above shows an example of a specific bearing that will eventually break down over time. It is the third bearing of the first dataset. The energy remains quite stable for around 700 cycles. Then the energy starts to increase rather fast. A huge peak can even be seen at the very end. This could show that the error in the bearing is causing a spike in energy for this bearing. In [15] it is stated that there is an overall consensus on this dataset on which point in time the bearing starts to break down. There are four phases stated in the paper. First the bearing is starting up, then the bearing is in a healthy state for a while. After this healthy state the phase of suspected error starts up until the actual fatal error occurs for the final phase. These phases have specific time windows when they start and when the dataset enters the next phase. To see if these phases show differences in energy the time windows were used to classify the data. In figure 19 the result of this will be shown in a scatter plot.
The scatterplot in figure 19 shows the result of classifying the energy of two features to the different phases. The green area shows the starting phase, the yellow shows the healthy area, the orange the suspected error and the red the fatal error. The points on the graph are the actual test data point colored to the window that they belong in. This shows that the test data falls into the classified areas from the training data. Some green point are located on the yellow area and also some orange point are located in the yellow area. This is logical because there is not a clear exact point when the motor changes phase. The picture does show however that a clear transition can be seen from one phase to the next by the increase of energy. At first one of the features stays quite steady until the suspected error phase when both energies increase until a big jump to the fatal error phase. The transitions are very well captured within this figure 19 which shows that the energy is a good feature for measuring the health index of the bearings.

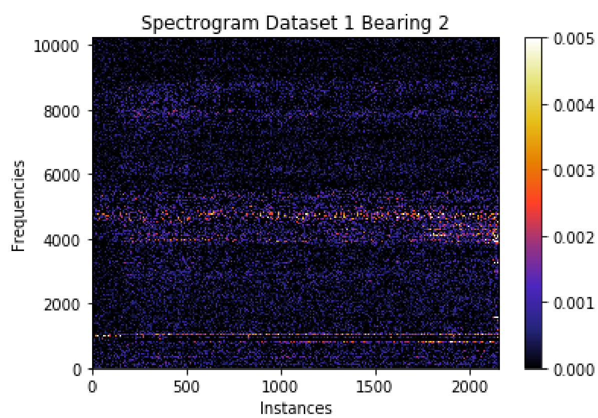
4.2.3 Spectrogram

The next step is to visualize the energy increase over the whole spectrum. This will show at what frequencies the energy increase occurs and if this increase happens at the frequencies where

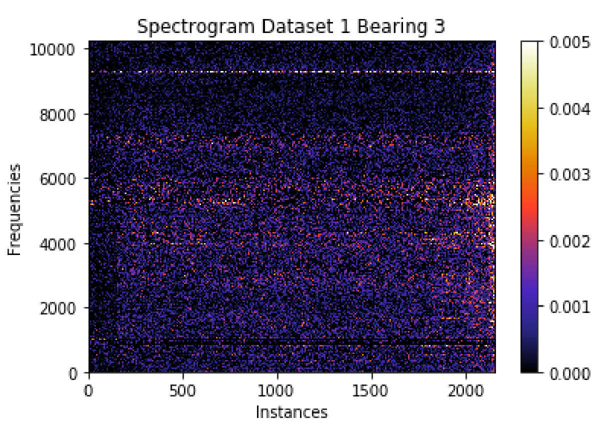
they are expected. The two errors that occur in these datasets are the Ball Pass Frequency Outer Race (BPFO) and the Ball Pass Frequency Inner Race (BPFI). These errors can be seen at specific frequencies and their harmonic orders according to [15]. The BPFI can be found at a frequency of 297 Hz while the BPFO can be found at a frequency of 236 Hz. These frequencies are the base frequencies at which these errors can be found, but as stated before, these errors can also be found at the different harmonic orders. These are the powers of these frequencies, so for 236 Hz it will be 472, 708, 944 and so on. This is why not only the base frequency but also the harmonic orders of the errors will be considered. First the full spectrum of all bearings in dataset one will be shown in figure 20. All frequencies until 10.000 Hz are shown here over time.
In figure 20 the following can be seen. Bearing one shows very steady frequencies. It has some frequencies which show higher energies than others but they stay quite consistent over the whole set. This can also be said about the second bearing. The frequencies stay quite steady. It gains a little more energy at the end but not over the whole spectrum. Bearing three on the other hand has higher energies over the whole spectrum. Also at 1500 cycles the overall energy seems to increase at a big range of the data. It is not an enormous spike but it is visible. For bearing four an enormous spike at the end is visible. This happens at a clear point at which the energy increases. From these spectrograms could be concluded that bearing three and four



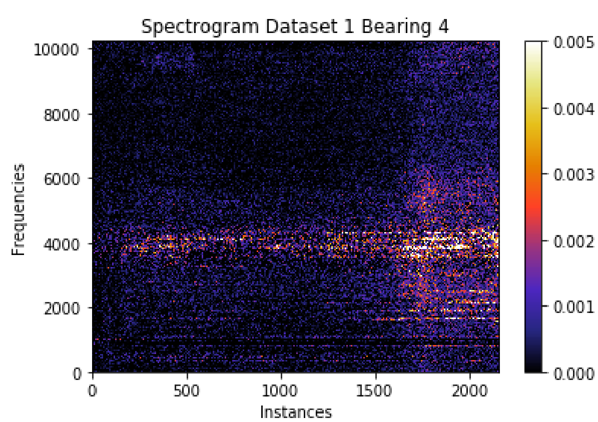
(a) Bearing 1



(b) Bearing 2



(c) Bearing 3



(d) Bearing 4

Figure 20: Spectrograms Dataset 1

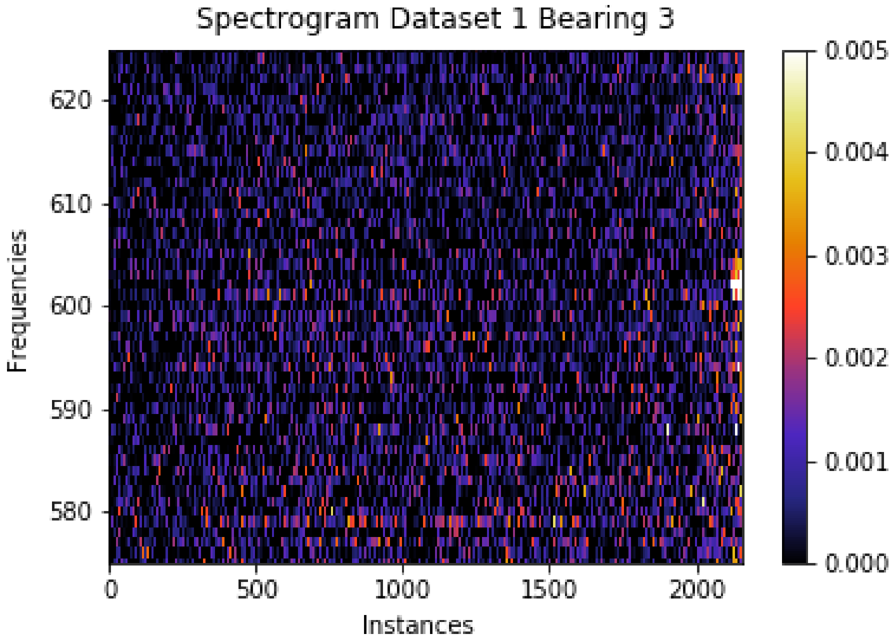
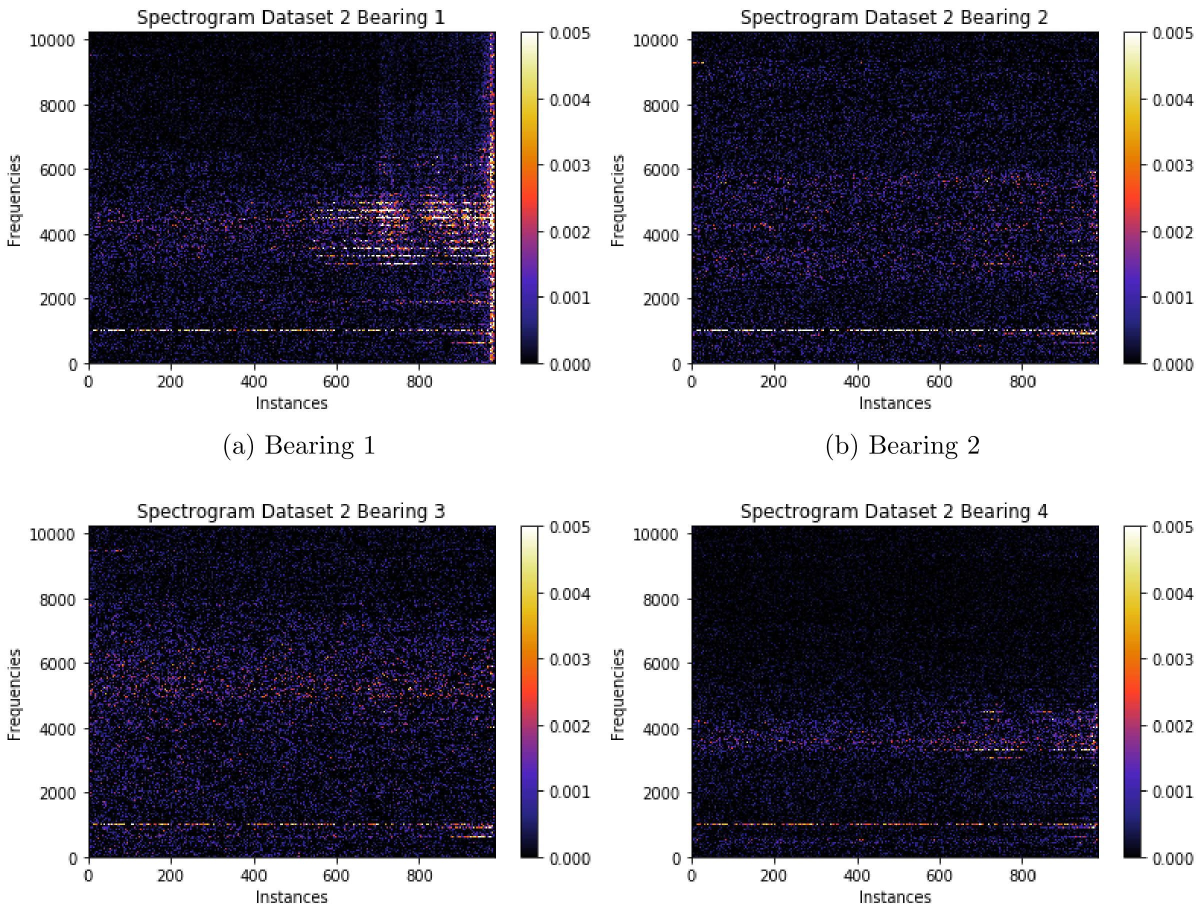


Figure 21: Spectrogram of the second harmonic order of the BPFI frequeny, IMS dataset 1

behave differently from the other two bearings. This is in line with what was expected. Bearing three breaks down at the end and after that bearing four also has a fatal error. Bearing three is the main bearing that breaks down in this dataset so now the spectrogram will be zoomed in to the specific frequencies of this error for bearing three. The second harmonic order, 594 Hz, is shown in figure 21.
Looking at figure 21 it is very difficult to see a big change at the specific frequency of 594 Hz. This could be due to the fact that the increase is present but not as considerable to pop up among the other frequencies. Other harmonic orders were also considered but these did not show any real visible increase in the spectrograms as well. The sum of these harmonic orders compared to the other bearings will be shown later on in the thesis. Now the next dataset will be shown in the same way as the first dataset. All four spectrograms together for the four bearings. The result can be seen in figure 22.
While the differences in the first dataset were not as big as expected. The second dataset shows exactly what was expected from the literature. Bearing one breaks down while the other bearings remain healthy. Bearing one shows a big increase in energy over time while the others show almost no increase except for the end of bearing four. Now to find the error that was found in this bearing the specific frequencies of the BPFO need to be visualized. This frequency is 236 and its harmonic orders. When the energy increases it can be said that this error occurs in this bearing. The third harmonic, frequency 708 Hz, is shown in figure 23.
Figure 23 shows what was expected from this interval. Around the frequency of 708 Hz an increase in energy occurs. The error also starts at the moment that was expected around half of the instances. This energy spike also occurs at different harmonic orders of the BPFO error. The figure 24 shows the eighth order, which is around 1888 Hz.
Also figure 24 shows the high increase around the expected frequency of 1888 Hz. Therefore it can be concluded that the BPFO is clearly present in this dataset and can be visualized using the energy.



(a) Bearing 1

(b) Bearing 2

(c) Bearing 3

(d) Bearing 4

Figure 22: Spectrograms Dataset 2

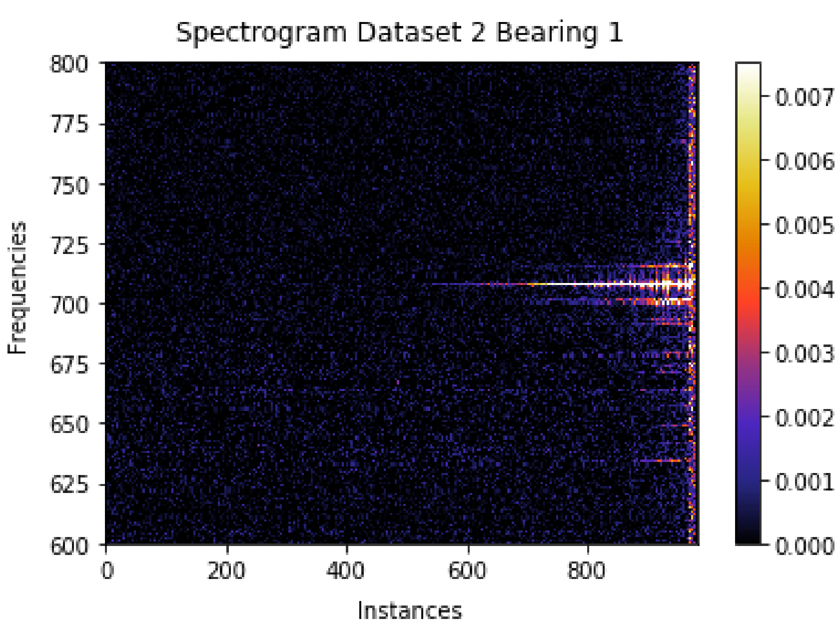


Figure 23: Third harmonic order of the BPFO frequeny, IMS dataset 2

Finally, the third dataset will be visualized using spectrograms. In [15] the BPFI was not found in the visualization. This is why it will be very interesting to see whether the visualization method proposed in this thesis can actually show the error. The error occurs in bearing three and is as stated before a BPFI error. All four spectrograms of the different bearings are shown in figure 25.
For this last dataset it is very difficult to see a lot of increase in the energy of all four bearings. For most of the instances the energy remains the same. Only at the very end an increase in energy can be observed. All four bearings show this but for bearing three it is the clearest. Which is also what was expected because the error occurs in bearing three. The reason the increase can be seen here but was not spotted in [15] could be because the amount of instances does not seem to be the same. While the dataset that was used for this thesis contains more than 6000 instances. The dataset of [15] only had a little over 4000 instances. This difference could be because new data was added later on. This could also explain that they could not see any difference because around 4000 instances there is no big difference visible.
To show that the error that seems to occur at the very end is in fact a BPFO, the harmonic orders of this error can again be considered in a smaller spectrogram. The base frequency is 236 Hz. For the following figure 26 the third harmonic order will be considered for the third bearing.
The spectrogram shows what was expected from the BPFO. The error seems to occur at the very end but a spike in energy around 708 Hz is clearly visible.
From all the spectrograms that were shown in this section it can be concluded that for dataset two and three the error can be shown using this method. The energy increases at the intervals where the increase was expected. For the first dataset it was not as visible as expected. The error did not show big spikes of energy while the others errors did. It will be interesting to see whether the error can still be observed in the health index despite not seeing it in the spectrogram.

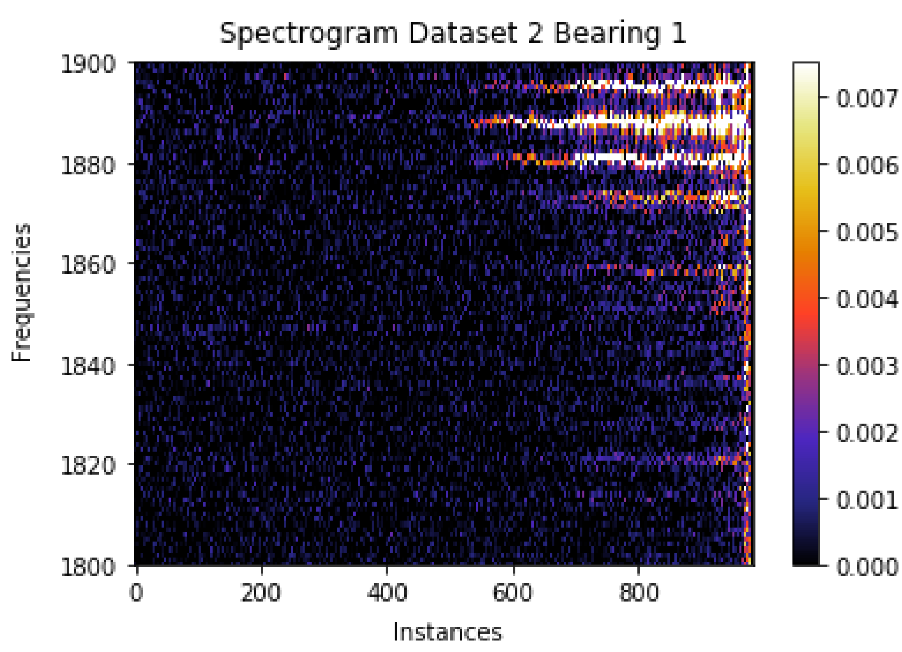
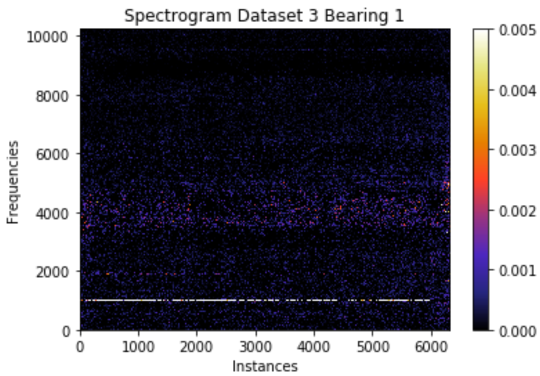
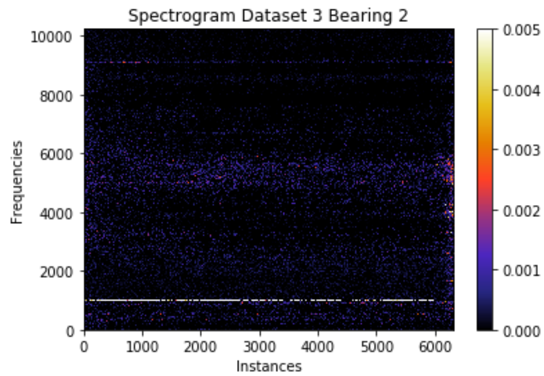


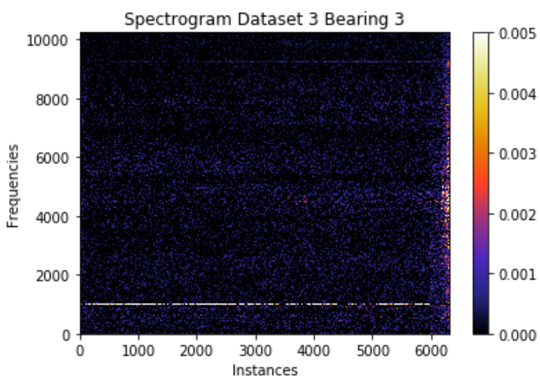
Figure 24: Eighth harmonic order of the BPFO frequeny, IMS dataset 2



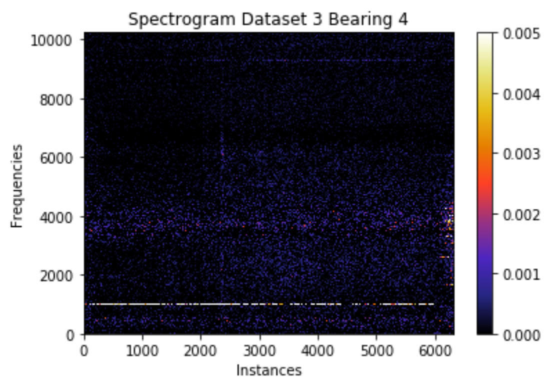
(a) Bearing 1



(b) Bearing 2



(c) Bearing 3



(d) Bearing 4

Figure 25: Spectrograms Dataset 3

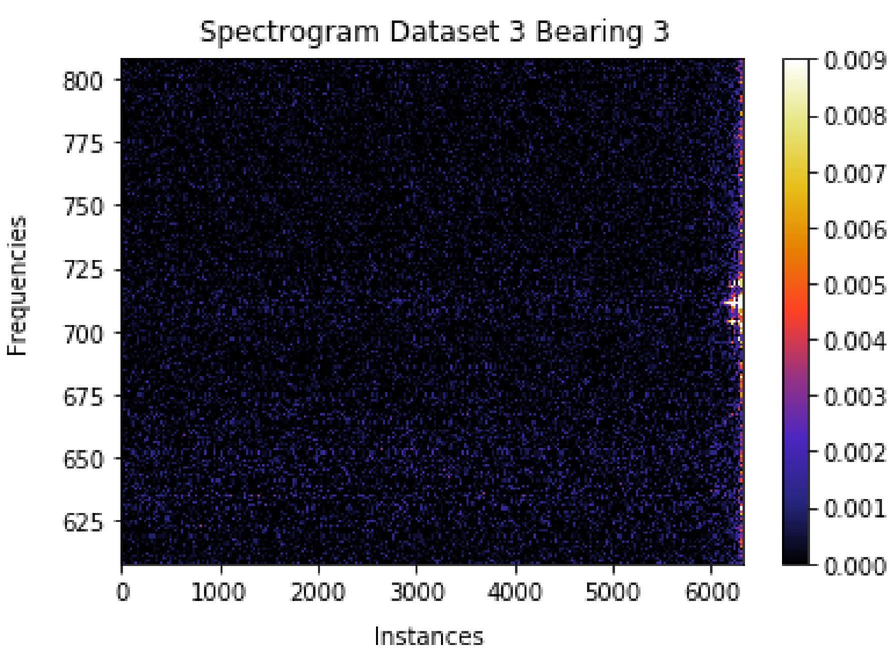


Figure 26: Third harmonic of the BPFI frequeny, dataset 3

4.2.4 Health and Failure Index

In the following section the actual failure rate will be calculated. The failure rate will show what bearing breaks down the fastest. A failure rate does not have a general way of calculation. The failure rate is based on a feature that can show whether a motor is failing or not. From the previous section it can be concluded that the energy is a good feature to show if a bearing is having an error or not. It also shows by the increase of energy how fast it breaks down. These two features of the energy make it a good candidate to use as the main calculation method for the failure rate.

The failure rate will now be calculated from the energy. This will be done by calculating the energy per instance for every bearing. Because we are looking for specific errors we will only consider the frequencies of these errors in the calculation. To make this more clear for the dataset one, only the energy around 297 Hz, 594 Hz, 891 Hz up until 11.880 Hz will be calculated and summed together. This comes down to the first 40 harmonic orders of the errors. The base frequency for dataset one is 297Hz and for dataset two and three it is 236 Hz. The sum of the energy of these harmonic orders can be seen in figures 27, 28 and 29 resulting in the failure rates.

The failure rates in the figures 27, 28 and 29 show some very interesting results. Dataset one shows one bearing that clearly fails the fastest compared to the other bearings. This bearing is the third bearing. This bearing also is the bearing that we expected here. Even though it wasn't possible to visualize that this error happens to this bearing, from the failure it actually is possible to conclude that bearing three breaks down the fastest considering the BPFI.

Figure 28 shows the opposite of what was expected from the spectrograms. The spectrograms show a clear difference in energy between the first bearing and the others. In the figure it can be seen that bearing three breaks down faster than bearing one up until the point where the

Figure 27: Failure Rate of the IMS Dataset 1, main Frequency of 297Hz

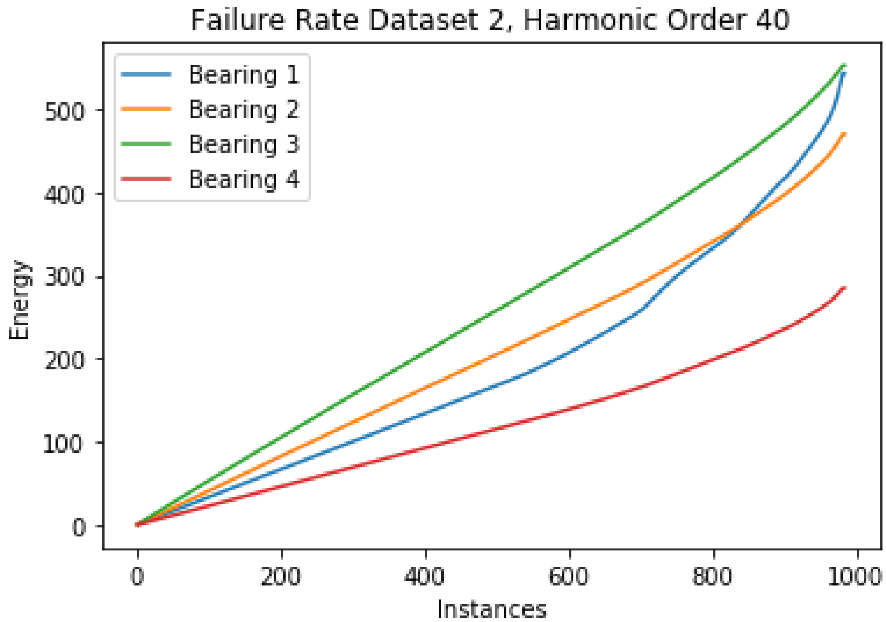


Figure 28: Failure Rate of the IMS Dataset 2, main Frequency of 236Hz

error seems to occur. This results in the fact that bearing three and one both end at more or less the same spot. Even though there is no evidence in other literature that bearing three breaks down as well. From this figure it could actually be stated that bearing three was actually breaking down as well. Unfortunately this is not verifiable because the tests stop when one of the bearings break down, in this case it was bearing one. So even though bearing one seems to break down the fastest at the end it is still a very interesting finding that bearing three could also be broken or is about to break as well.

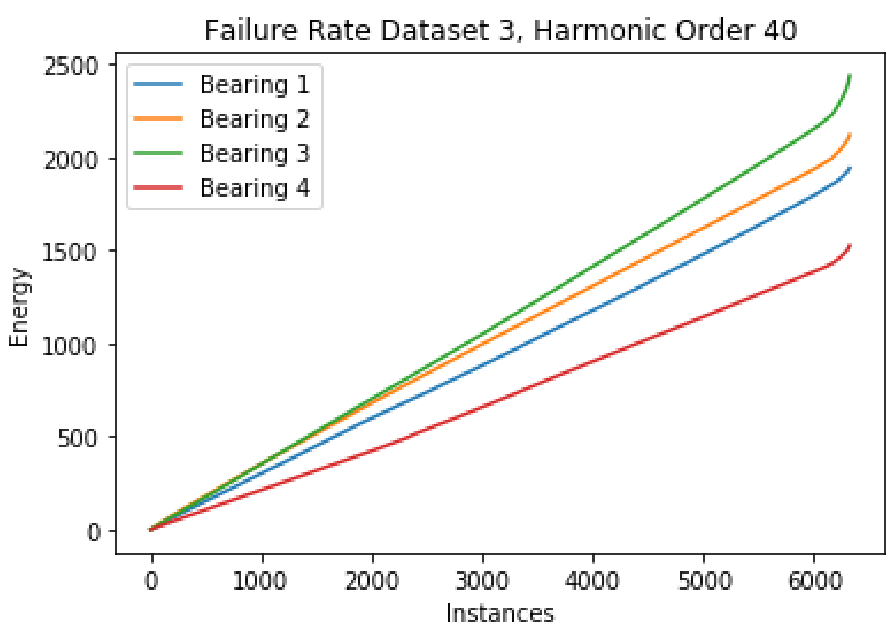


Figure 29: Failure Rate of the IMS Dataset 3, main Frequency of 236Hz

Figure 29 shows what was expected from this dataset. Overall all bearings stay quite steady over the whole dataset up until the very end. Bearing three, even though slightly, seems to break down the fastest. The interesting part for this dataset is the energy spike at the very end for all bearings. This was also visible in the spectrograms. This could have different causes. The first possibility is that when bearing three breaks down the energy increase in that bearing increases the energy for all other bearings as well. The bearings are connected in the test set up, so this could be a possibility. The other option is that something goes wrong when the test ends. The ending condition of the datasets is not totally clear. So the sudden huge increase could also be a result of human interference at the very end, but this is hard to tell because there is no real data on how the tests ended specifically.

From the figures discussed in this section can be concluded that the failure rate that is calculated from the energy feature works well on the given dataset set. The failures that were expected also show in these figures. An interesting observation is that in all cases bearing three is the bearing fastest increasing failure rate. For dataset two bearing one eventually has its failure rate increase faster but after the start bearing three increases faster at first. This is why it might not be a coincidence that bearing three breaks down in both dataset one and three. It could be that the way the test was set up bearing three has a specific conditions which makes its health decrease faster than the others. These specific results could be interesting to explore in more depth in future work.

The next step is to calculate the health index from this failure rate. The health index is an index that starts at one for fully healthy up until zero where the motor is broken. The faster this line will go down to zero the faster the motor breaks down. We calculate this index from the failure rate by first normalizing the data between zero and one. Now the failure rate needs to be normalized so that it will go from zero to one. For the health index the opposite is required. To achieve this all values in the graph will be reversed. This means that zero becomes one, one becomes zero and everything in between will also be inverted. This results in the health

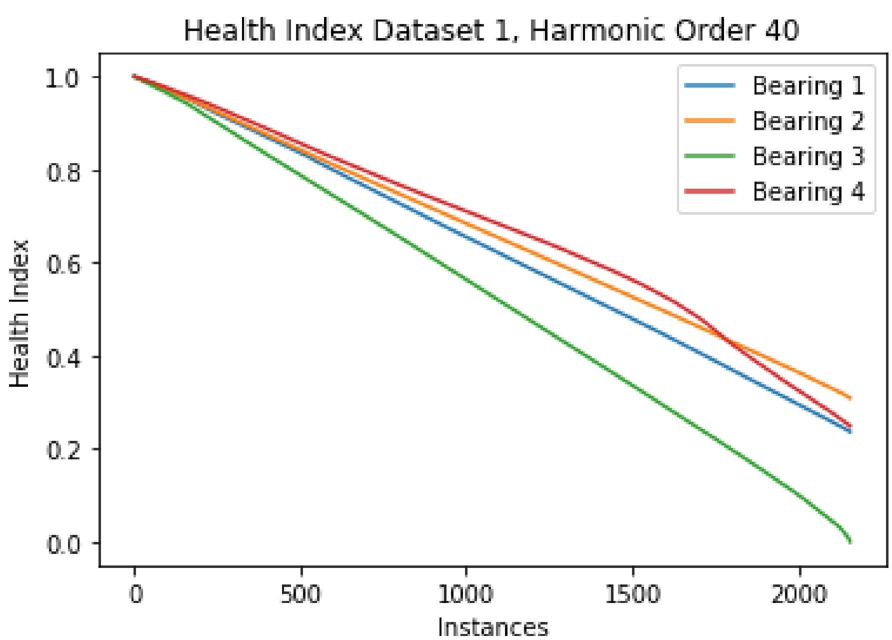


Figure 30: Health Index of the IMS Dataset 1

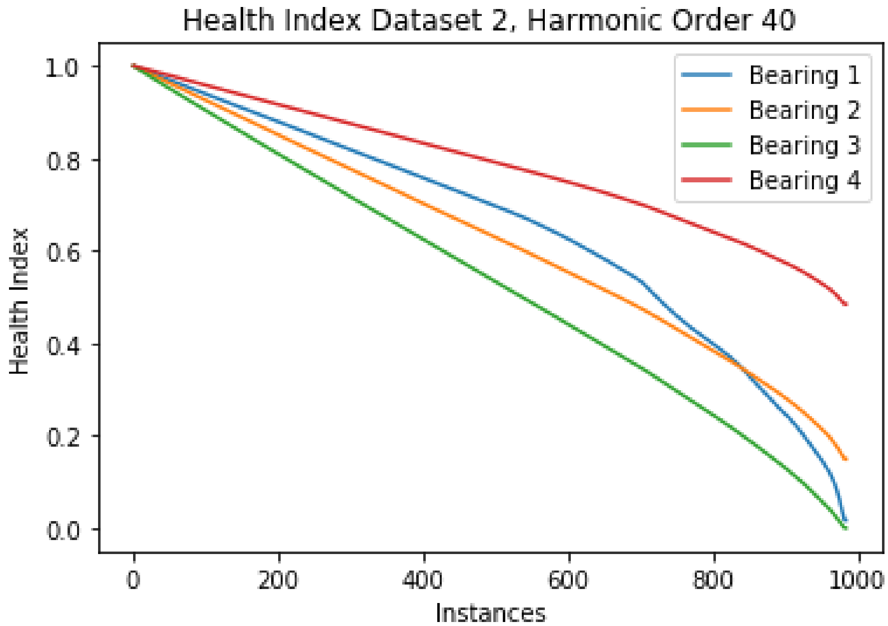


Figure 31: Health Index of the IMS Dataset 2

index graphs. These graphs can be seen in figures 30, 31 and 32. All three datasets are shown in order.
The health indexes in figures 30, 31 and 32 show the same results as was seen in the figures from the failure rate. It can be seen that for two of the datasets, dataset one and three, that the bearing that decreases the fastest is actually the one that breaks down.The second dataset shows that at the end of the test both bearing one and three reach zero. It is also visible that bearing one starts to break down very fast compared to the other bearings. This indicates that,

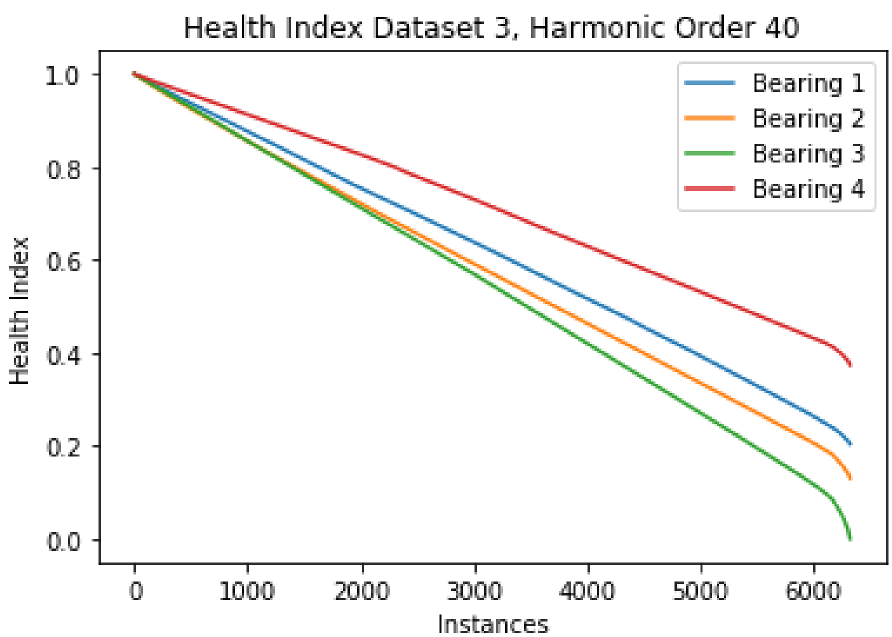


Figure 32: Health Index of the IMS Dataset 3

while bearing one's health started off decreasing slower than bearing three's health, bearing one's health starts to decrease a lot faster later on and thus could show a more severe error than bearing three. It is still interesting to see the health index of bearing three is zero at the end of the test even though there is no mention of this in other literature. Unfortunately whether bearing three was about to break is not verifiable because the tests end when one of the bearings is broken, in this case bearing three. Overall the results seem very promising because of the fact that the expected behaviour also shows in the graphs. This leads to the final step, the calculation of the remaining useful lifetime for the other bearings.

4.2.5 Remaining Useful Lifetime

The remaining useful lifetime will be calculated on top of the health index results. The remaining useful lifetime shows how many cycles the remaining bearings could have stayed healthy. This is done by predicting the path of the health index over time. This means that the remaining useful lifetime will show how the health index would further behave when the test would not have been stopped. This prediction is done using the fbprophet [32] package. This package takes time series data and predicts how it will behave over time. The time series data that fbprophet receives will be the health index data. To clearly see how much cycles the bearings have left as remaining useful lifetime the graph will be transposed. The graph will be transposed so that the health index data will go up until zero on the x axis. All data after the zero on the x axis will therefore be predicted data. When the predicted lines reach zero the amount of cycles that were predicted as the remaining useful lifetime can be read. Fbprophet is, as stated before, a prediction package. This means that the predictions always have a range in which they fall. This is visualized in the graphs by the colored areas around the lines. These are the minimal and maximal range of the predicted values. Just as all the other calculations, the remaining useful lifetime was calculated by looking at the first 40 harmonic intervals of the base frequency of the errors. So for dataset one a frequency of 297 Hz and its powers, for dataset two and three a frequency of 236 Hz.

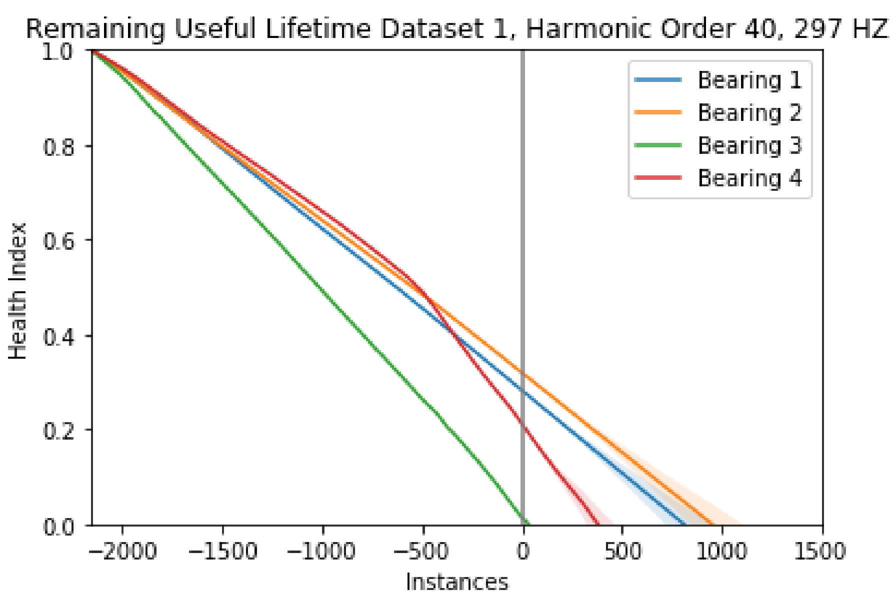


Figure 33: Remaining Useful Lifetime of IMS Dataset 1

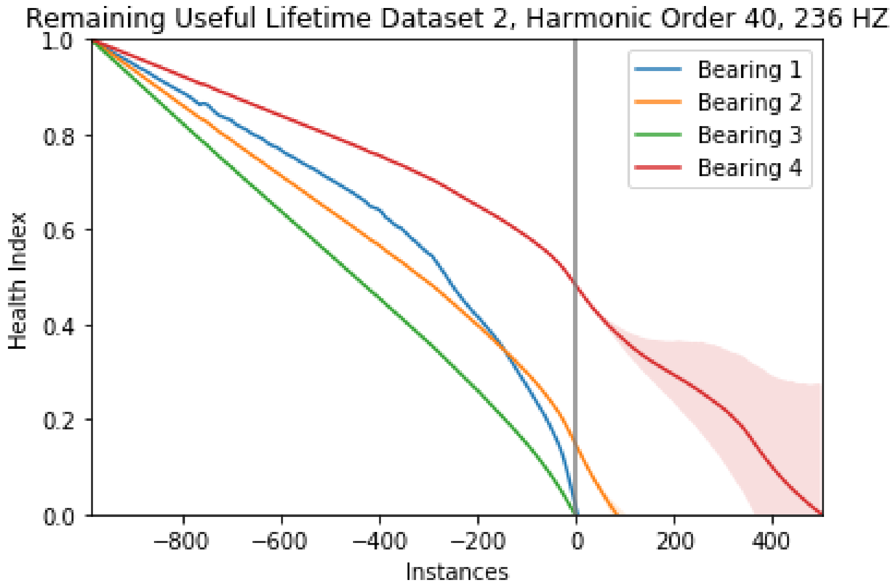


Figure 34: Remaining Useful Lifetime of IMS Dataset 2

The final results can be seen in figures 33, 34 and 35. The vertical line shows the split between health index data and the predicted remaining useful lifetime data.
Figures 33, 34 and 35 show the results of the predicted remaining useful lifetime. The results mostly show quite predictable results. For example figure 33 shows the lines moving on consistently until they finally reach zero. Also the range of predictions do not show odd behaviour. For the other two graphs there are some interesting results to be discussed. The most interesting result is the predicted lifetime of bearing four of dataset two. This shows a huge range in the predicted values. The reasons this could happen is because of patterns in the health index data that are not clearly visible. Even though these patterns can not be seen on the

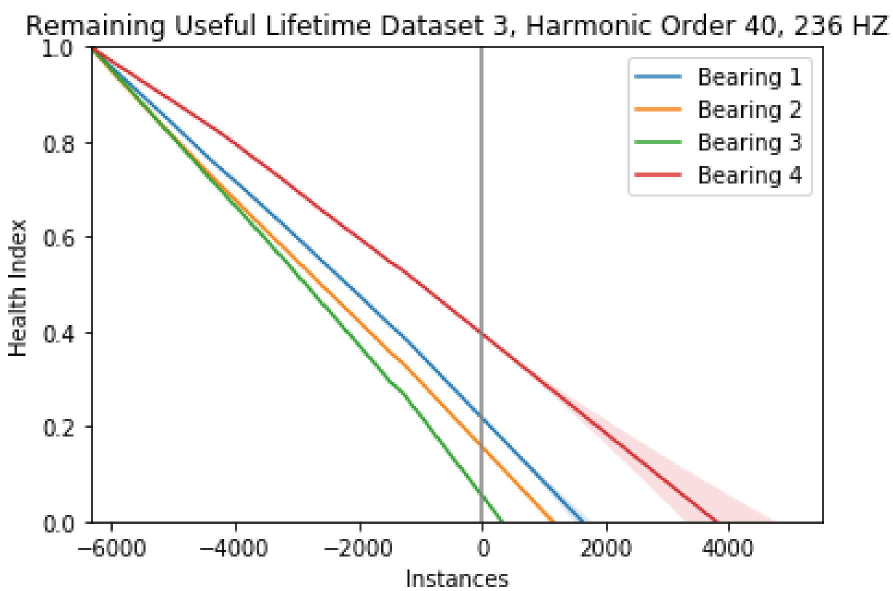


Figure 35: Remaining Useful Lifetime of IMS Dataset 3

graph, fbprophet might see them in the data. This could result in the huge prediction range. The other result that should be discussed is the fact that bearing three does not end on zero. The explanation for this is that fbprophet probably does not consider the very ending of the health index as important as the overall trend of the dataset. If these sudden changes at the very end would have been considered, all bearing would have only lasted a few more cycles. Because the ending conditions of the tests are not clear it is better to predict the remaining useful lifetime on the overall trend of the dataset and thus this figure shows that with that knowledge it might have had a few more cycles before a fatal error.
This thesis used the harmonic order of 40 for the different calculations. The reason for picking this number is that from 40 on and higher the difference in remaining useful lifetime does not change much anymore. The lower orders show varying results but from 40 and up the number stay quite similar. This can also be seen in the following tables. The tables show the harmonic order and what the remaining useful lifetime would have been for the healthy bearings if this harmonic order would have been chosen.

H. O.	Bearing 1			Bearing 2			Bearing 4		
	Mean	Lower	Upper	Mean	Lower	Upper	Mean	Lower	Upper
5	859	744	1046	620	582	671	961	784	1390
10	787	668	1008	1122	1015	1279	665	549	895
15	405	367	461	1103	919	1498	-1	-2	0
20	365	333	410	760	663	897	182	159	220
25	668	601	785	1098	938	1333	408	346	530
30	740	660	873	954	844	1130	460	388	603
35	809	710	945	951	846	1110	378	326	471
40	809	715	953	951	841	1107	378	324	462
45	809	721	953	951	850	1112	378	323	470
50	809	719	969	951	850	1108	378	326	498

Table 11: Dataset 1

From these tables we can conclude that from the harmonic order of 40 and up the predicted remaining useful lifetime does not vary a lot anymore. This is why 40 was a good amount to use as the basis for this thesis.
Looking back at the complete calculation of the remaining useful lifetime it can be concluded

H. O.	Bearing 2			Bearing 3			Bearing 4		
	Mean	Lower	Upper	Mean	Lower	Upper	Mean	Lower	Upper
5	99	79	319	-4	-4	-4	427	289	572
10	50	45	63	-4	-4	-3	325	213	601
15	15	15	16	-4	-4	-3	298	192	614
20	107	86	316	100	85	137	356	251	610
25	122	98	313	69	61	80	438	319	622
30	113	92	196	24	23	25	468	347	638
35	90	76	127	-2	-2	-2	474	349	647
40	82	71	100	-4	-4	-3	498	362	623
45	87	75	116	-4	-4	-3	498	362	629
50	87	75	108	-4	-4	-3	498	362	627

Table 12: Dataset 2

H. O.	Bearing 1			Bearing 2			Bearing 4		
	Mean	Lower	Upper	Mean	Lower	Upper	Mean	Lower	Upper
5	3399	3014	3976	2662	2400	3026	5801	4953	7469
10	6822	5772	9176	2603	2338	2966	8706	6965	9982
15	2441	2210	2785	1886	1741	2042	3856	3295	4786
20	968	910	1036	1794	1643	2008	2440	2225	2723
25	1020	962	1094	1219	1139	1309	2863	2528	3391
30	1082	1017	1168	968	916	1025	3094	2726	3645
35	1401	1306	1527	1043	983	1122	3527	3083	4146
40	1622	1489	1770	1144	1079	1223	3788	3279	4747
45	1668	1546	1823	1124	1055	1202	3714	3288	4326
50	1668	1545	1834	1124	1054	1198	3714	3295	4314

Table 13: Dataset 3

that the approach taken in this thesis is very promising. The energy proves to be a very helpful feature both showing the current health index as well as calculating the remaining useful lifetime of these bearings.

5 Conclusion

5.1 Classification

The thesis will be summarized by answering the research questions stated in section 1.5

5.1.1 Summary

How can raw data from the current of a motor be used to classify whether a motor is broken and what error occurred?

In this thesis, we successfully detected the type of errors of electrical motor as well as estimated the remaining useful lifetime of bearings. Both solutions came from the calculated energy that was retrieved from the Fast Fourier transformation of the raw data. This energy was taken from specific intervals around the main peaks and between them.
The classification resulted in using the energy of specifically chosen windows and classify them using algorithms like logistic regression and support vector machines. In combination with an increased class weight this resulted in a minimal amount of false negatives and an acceptable amount of false positives. By changing the class weight the amount can be varied if that would be more preferable. By showing the results of the motors combined it was shown that this method is not only fit for one specific type of motor.

5.1.2 Discussion and Future Work

For this specific research we only used the main frequency of 50 Hz. The signal also had different main frequencies like 25 Hz but there wasn't an easy way to generalize this data so that both main frequencies could be used.
The future work that can be done on this subject is most of all finding a way to generalize the data so that the main frequency is not a factor in the classification. This would be a major improvement to this method. A possible suggestion for tackling this issue is to generalize the raw data to a specific chosen frequency. This can only be done if the main frequency is known beforehand. Another approach could be translating the Fast fourier transformation to a standard frequency. The result could be wider or thinner peaks. This might not be what you would want because of the energy feature. The data should therefore also be normalized. A final approach could be to filter out the main frequency from either the raw data or the FFT. Only data that is not related to the main frequency will then be shown in the FFT. This way the same error frequencies can be analyzed on the dataset. Because of these reasons this approach seems the most promising for future work. In [33] multiple approaches to this filtering are proposed together with their advantages and disadvantages. Furthermore it would also be interesting to try this method on other datasets to see how well it performs.

5.2 Remaining Useful Lifetime

5.2.1 Summary

How can vibration data from bearing be used for the calculation of the remaining useful lifetime?

In this thesis the remaining useful lifetime was estimated by using the energy that was calculated from the fast Fourier transformation. In spectrograms was shown that the different harmonic orders of the specific errors could be used to calculate a failure rate. This failure rate was then calculated by taking the summation of all the energy around these harmonic order frequencies. This resulted in the failure index which could be translated to the health index. The health index could then be used to forecast the further remaining useful lifetime until it hit a specific boundary. By doing this for three distinct dataset it was shown that this method can be used on different set ups and motors.

5.2.2 Discussion and Future work

The main discussion point with this research is also the main frequency issue. The dataset used for this research had only one specific main frequency which was good to have a start with but it at this point it is not possible to work with other main frequencies yet or with frequency changes. This is why this would be the most interesting for future work as well. This could be a huge benefit if this method can be translated to a dataset with a varying main frequency. The methods mentioned in section 5.1.2 could therefore be looked into. Also the fact that bearing three's health index decreases fast in all three datasets could be interesting to look into in future work.

6 Acknowledgement

In the process of writing this thesis I have received help to achieve the final result. In this section I would like to thank them for their support.

First and foremost I would like to thank Dr. D. V. Nguyen who has supervised me throughout the whole thesis. With his help and supervision on every part he helped me achieve all the promising results in this thesis. I had a very good time working with Dr. D. V. Nguyen and it was nice to have someone to spar with in the weekly meetings. His feedback from the start up until the final result really helped me write this thesis to the best of my ability.

Secondly I would like to thank Semiotic Labs for the use of their dataset and the feedback that they have given me in several meetings.

Finally I would like to thank Prof. dr. T. H. W. Bäck for helping me find this project and giving me feedback on the final results.

7 Bibliography

References

[1] D. Siyambalapitiya and P. Mclaren, "Reliability improvement and economic benefits of on-line monitoring systems for large induction machines," *IEEE Trans. Ind. Appl.*, vol. 26, pp. 1018 – 1025, 1990.

[2] S. Kumar, D. Mukherjee, P. K. Guchhait, R. Banerjee, A. K. Srivastava, D. N. Vishwakarma, and R. K. Saket, "A Comprehensive Review of Condition Based Prognostic Maintenance (CBPM) for Induction Motor," *IEEE Access*, vol. 7, pp. 90 690–90 704, 2019. [Online]. Available: https://ieeexplore.ieee.org/document/8754739/

[3] P. Allbrecht, J. Appiarius, R. McCoy, and E. Owen, "Assessment of the reliability of motors in utility applications—updated," *IEEE Trans Energy Convers*, pp. 39–46, 1986.

[4] V. V. Thomas, K. Vasudevan, and V. J. Kumar, "Implementation of online air-gap torque monitor for detection of squirrel cage rotor fault using tms320c31," *Int. Conf. on Power Electron., Machi. Drives,*, pp. 128–132, 2002.

[5] P. Vas, *Parameter estimation, condition monitoring and diagnosis of electrical machines*. Oxford: Clarendon Press, Oxford, 1993.

[6] W. Deleroi, "Broken bars in squirrel cage rotor of an induction motor- part i: description by superimposed fault currents," *Arch Elektrotech*, pp. 91–99, 1984.

[7] Scheme-3-phases-inductive-motor-design.jpg (500281). [Online]. Available: https://eltra-trade.com/files/uploads/news/Scheme-3-phases-inductive-motor-design.jpg

[8] Failure modes for bearings. [Online]. Available: https://www.emersonbearing.com/bearing-failure-modes/

[9] A. Cubillo, S. Perinpanayagam, and M. Esperon Miguez, "A review of physics-based models in prognostics: Application to gears and bearings of rotating machinery," *Advances in Mechanical Engineering*, vol. 8, 08 2016.

[10] S. Mohamed, T. Tettey, and T. Marwala, "An extension neural network and genetic algorithm for bearing fault classification," in *The 2006 IEEE International Joint Conference on Neural Network Proceedings*, 2006, pp. 3942–3948.

[11] A. Dineva, A. Mosavi, M. Gyimesi, and I. Vajda, "Multi-label classification for fault diagnosis of rotating electrical machines." [Online]. Available: http://www.preprints.org/manuscript/201908.0029/v1

[12] A. K. Mahamad, S. Saon, and T. Hiyama, "Predicting remaining useful life of rotating machinery based artificial neural network," vol. 60, no. 4, pp. 1078–1087. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0898122110002555

[13] A. Qin, Q. Zhang, Q. Hu, G. Sun, J. He, and S. Lin, "Remaining useful life prediction for rotating machinery based on optimal degradation indicator," *Shock and Vibration*, vol. 2017, pp. 1–12, 03 2017.

[14] A. Mosallam, "Remaining useful life estimation of critical components based on bayesian approaches." p. 146.

[15] W. Gousseau, J. Antoni, F. Girardin, U. Lyon, and J. Griffaton, "Analysis of the rolling element bearing data set of the center for intelligent maintenance systems of the university of cincinnati," p. 14.

[16] Semiotic labs: Eliminate all unplanned downtime. [Online]. Available: https://www.semioticlabs.com

[17] Prognostics center of excellence - data repository. [Online]. Available: https://ti.arc.nasa.gov/tech/dash/groups/pcoe/prognostic-data-repository/

[18] Schema-of-the-test-rig-at-IMS-of-university-of-cincinnati-by-courtesy-of-ref-18.png (850579). [Online]. Available: https://www.researchgate.net/profile/A_Lakis/publication/270776477/figure/fig6/AS:392145216262145@1470506113495/Schema-of-the-test-rig-at-IMS-of-University-of-Cincinnati-by-courtesy-of-Ref-18.png

[19] R. Verdelli. Data science process: Data exploration. [Online]. Available: https://www.techedgegroup.com/blog/data-science-process-data-exploration

[20] Data preprocessing in machine learning. [Online]. Available: https://serokell.io/blog/data-preprocessing

[21] U. Oberst, "The fast fourier transform," vol. 46, no. 2, pp. 496–540. [Online]. Available: http://epubs.siam.org/doi/10.1137/060658242

[22] C. Maklin. Fast fourier transform. [Online]. Available: https://towardsdatascience.com/fast-fourier-transform-937926e591cb

[23] P. P. Ippolito. Feature extraction techniques. [Online]. Available: https://towardsdatascience.com/feature-extraction-techniques-d619b56e31be

[24] S. Haimoura. Resampling methods for unbalanced datasets — fraudulent transactions. [Online]. Available: https://towardsdatascience.com/https-towardsdatascience-com-resampling-methods-for-unbalanced-datasets-5b565d0a247d

[25] scikit-learn: machine learning in python — scikit-learn 0.23.2 documentation. [Online]. Available: https://scikit-learn.org/stable/

[26] P. Gupta. Decision trees in machine learning. [Online]. Available: https://towardsdatascience.com/decision-trees-in-machine-learning-641b9c4e8052

[27] T. Yiu. Understanding random forest. [Online]. Available: https://towardsdatascience.com/understanding-random-forest-58381e0602d2

[28] S. Swaminathan. Logistic regression — detailed overview. [Online]. Available: https://towardsdatascience.com/logistic-regression-detailed-overview-46c4da4303bc

[29] R. Gandhi. Support vector machine — introduction to machine learning algorithms. [Online]. Available: https://towardsdatascience.com/support-vector-machine-introduction-to-machine-learning-algorithms-934a444fca47

[30] Legacy discrete fourier transforms (scipy.fftpack) — SciPy v1.5.4 reference guide. [Online]. Available: https://docs.scipy.org/doc/scipy/reference/fftpack.html

[31] Practical tips for class imbalance in binary classification | by zichen wang | towards data science. [Online]. Available: https://towardsdatascience.com/practical-tips-for-class-imbalance-in-binary-classification-6ee29bcdb8a7

[32] Quick start. [Online]. Available: http://facebook.github.io/prophet/docs/quick_start.html

[33] A. de Cheveigné and I. Nelken, "Filters: When, why, and how (not) to use them," vol. 102, no. 2, pp. 280–293. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0896627319301746