

ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL

**DATA-DRIVEN CONDITION MONITORING AND
FAULT DIAGNOSIS OF VFD-FED
INDUCTION MOTORS**

M.Sc. THESIS

Alper SENEM

Department of Mechatronics Engineering

Mechatronics Engineering Programme

JUNE 2021

ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL

**DATA-DRIVEN CONDITION MONITORING AND
FAULT DIAGNOSIS OF VFD-FED
INDUCTION MOTORS**

M.Sc. THESIS

**Alper SENEM
(518181003)**

Department of Mechatronics Engineering

Mechatronics Engineering Programme

Thesis Advisor: Prof. Dr. Şeniz ERTUĞRUL

JUNE 2021

İSTANBUL TEKNİK ÜNİVERSİTESİ ★ LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

**DEĞİŞKEN FREKANSLI SÜRÜCÜ
İLE BESLENEN ASENKRON MOTORLarda
VERİ ODAKLI DURUM İZLEME VE ARIZA TANILAMA**

YÜKSEK LİSANS TEZİ

**Alper SENEM
(518181003)**

Mekatronik Mühendisliği Anabilim Dalı

Mekatronik Mühendisliği Programı

Tez Danışmanı: Prof. Dr. Şeniz ERTUĞRUL

HAZİRAN 2021

Alper SENEM, a M.Sc. student of ITU Graduate School student ID 518181003, successfully defended the thesis entitled “DATA-DRIVEN CONDITION MONITORING AND FAULT DIAGNOSIS OF VFD-FED INDUCTION MOTORS”, which he/she prepared after fulfilling the requirements specified in the associated legislations, before the jury whose signatures are below.

Thesis Advisor : **Prof. Dr. Şeniz ERTUĞRUL**
Istanbul Technical University

Jury Members : **Prof. Dr. Mustafa BAĞRIYANIK**
Istanbul Technical University

Prof. Dr. Levent Eren
Izmir University of Economics

Date of Submission : 11 June 2021
Date of Defense : 6 July 2021

To my grandparents, in loving memory.

FOREWORD

First of all, I would like to thank my advisor, Prof. Dr. Şeniz Ertuğrul, for her guidance and support since my undergraduate years. Thank you Prof. Dr. Mustafa Bağriyanık and Prof. Dr. Levent Eren for agreeing to serve on my committee. I would also like to express my sincere gratitude to Dr. Murat Imeryüz for teaching me about induction motors and for his valuable contributions.

I would like to thank my friend and colleague Burak Gökçen for his unwavering support from the very first day, and Sarper Öztürk, who made up for what I lacked with his knowledge of induction motors and conducted long and boring experiments with fun. I would also like to thank Wat Motor R&D Manager Barış Tuğrul Ertuğrul for his support and belief in me, and my team leader Sinan Unan for his understanding throughout my thesis.

I cannot thank enough to my beloved family and my dear fiance Damla Özyazgan, who have not spared their faith and support for me, for being there for me whenever I need it. If it were not for them, I would not be here.

Last but not least, I am grateful to the technicians at the Wat Motor General Purpose Industrial Motor Laboratory for their crafty handiwork and effort.

June 2021

Alper SENEM
(Mechanical Engineer)

TABLE OF CONTENTS

	<u>Page</u>
FOREWORD.....	ix
TABLE OF CONTENTS.....	xi
ABBREVIATIONS	xiii
SYMBOLS	xv
LIST OF TABLES	xvii
LIST OF FIGURES	xix
SUMMARY	xxi
ÖZET	xxiii
1. INTRODUCTION	1
1.1 Overview	1
1.2 Objectives of Research	3
1.3 Organization of Thesis.....	3
2. CONDITION MONITORING OF INDUCTION MOTORS: BACK-GROUND	5
2.1 Introduction of Induction Motors	5
2.1.1 Principle of operation.....	5
2.1.2 VFD-fed induction motors.....	6
2.1.3 Need for condition monitoring.....	7
2.1.4 Maintenance strategies.....	8
2.2 Induction Motor Fault Types	10
2.2.1 Bearing related faults	11
2.2.2 Stator related faults	12
2.2.3 Rotor related faults.....	14
2.3 Condition Monitoring Techniques.....	15
2.3.1 Temperature monitoring.....	16
2.3.2 Vibration monitoring.....	16
2.3.3 Motor current monitoring	17
2.4 Signal Processing Techniques	17
2.4.1 Time domain based signal analysis.....	18
2.4.2 Statistical analysis.....	18
2.4.2.1 Mean	18
2.4.2.2 Median	19
2.4.2.3 Root mean square	19
2.4.2.4 Standard deviation	19
2.4.2.5 Kurtosis.....	19
2.4.2.6 Skewness.....	20
2.4.3 Frequency based signal analysis	20
2.4.3.1 Shannon-Nyquist sampling theory	20

2.4.3.2 Fast Fourier transform	21
2.4.3.3 Power spectral density estimation.....	22
2.5 Data-driven fault diagnosis techniques.....	24
2.5.1 Machine learning methods.....	24
2.5.1.1 Naive Bayes	24
2.5.1.2 k-Nearest neighbors	25
2.5.1.3 Support vector machines.....	25
2.5.1.4 Multi layer perceptron	25
2.5.1.5 Ensemble learning	26
Random forest.....	26
XGBoost	26
2.5.2 Deep learning methods	27
2.5.2.1 1D-Convolutional neural networks.....	27
2.5.2.2 Long-Short term memory networks.....	28
2.6 Performance evaluation	28
2.6.1 Precision and recall	29
2.6.2 Accuracy	29
2.6.3 F-measure.....	30
2.6.4 Cohen's kappa.....	30
2.6.5 Area under the curve	31
3. EXPERIMENTAL SETUP AND METHODOLOGY	33
4. FAULT DIAGNOSIS METHODOLOGY	49
4.1 Machine learning analysis	49
4.1.1 Time domain statistical analysis	49
4.1.2 Frequency domain statistical analysis.....	52
4.1.3 Statistical analysis on characteristic frequencies	53
4.1.4 Discussion	53
4.2 Deep learning analysis.....	56
4.3 Discussion.....	58
5. CONCLUSIONS AND RECOMMENDATIONS.....	61
REFERENCES.....	65
APPENDICES	73
APPENDIX A.1	75
APPENDIX B.1.....	77
APPENDIX C.1.....	79
CURRICULUM VITAE.....	83

ABBREVIATIONS

VFD	: Variable Frequency Drive
AFD	: Adjustable Frequency Drive
HVAC	: Heating, Ventilation and Air Conditioning
AC	: Alternating Current
DC	: Direct Current
PCA	: Principle Component Analysis
PWM	: Pulse Width Modulation
V/f	: Voltage over Frequency
FOC	: Field Oriented Control
DTC	: Direct Torque Control
DFT	: Discrete Fourier Transform
FFT	: Fast Fourier Transform
PSD	: Power Spectral Density
SNR	: Signal-to-Noise Ratio
RMS	: Root Mean Square
STD	: Standard Deviation
SVM	: Support Vector Machines
kNN	: k-Nearest Neighbor
MLP	: Multi-Layer Perceptron
XGBOOST	: Extreme Gradient Boosting
CNN	: Convolutional Neural Networks
RNN	: Recurrent Neural Networks
LSTM	: Long-Short Term Memory
ROC	: Receiver Operating Characteristic
AUC	: Area Under the ROC Curve
TP	: True Positive
FP	: False Positive
FN	: False Negative
TN	: True Negative
t-SNE	: T-distributed Stochastic Neighbor Embedding
GPU	: Graphical Processing Unit
TPU	: Tensor Processing Unit
FPGA	: Field Programmable Gate Array

SYMBOLS

f_e	: Electrical supply frequency
f_v	: Rotational speed frequency of the rotor
m	: Harmonic number
f_{bng}	: Current component frequency due to air gap changes
f_{st}	: Principle slot harmonics frequencies
s	: Slip
p	: Number of pole pairs of the motor
f_{brb}	: Broken rotor bar sideband frequencies
f_s	: Sampling frequency
T_s	: Sampling period
Δf	: Spectral resolution
N	: Number of data samples
\bar{Y}	: Mean
\tilde{Y}	: Median
c_i	: Reference class
C	: Total number of classes
$p(c_i)$: Prevalence of the reference class
$p(c_i)$: Area under the class reference ROC curve for c_i
p_k	: Number of instances that class k was predicted
t_k	: Number of instances that class k actually presents
d	: Total number of correctly predicted samples

LIST OF TABLES

	<u>Page</u>
Table 2.1 : Distribution of induction motor faults by component (%).....	10
Table 2.2 : Actual distribution of multiple failures (%) [1].....	10
Table 3.1 : Equipments used in experimental-setup.	33
Table 3.2 : Nominal values of WAT Motor 3-phase induction motor.....	34
Table 3.3 : Brief information about experimental conditions.....	34
Table 3.4 : Input parameters for estimating Welch's PSD.....	42
Table 4.1 : Performance metrics for methods and classifiers with 5-fold cross-validation.....	55
Table 4.2 : 1D- CNN structure.....	57
Table 4.3 : LSTM structure.....	58
Table 4.4 : Performance metrics for deep learning architectures.....	58
Table B.1 : Naive Bayes algorithm parameters.	77
Table B.2 : Multi-Layer Perceptron algorithm parameters.....	77
Table B.3 : Random Forest algorithm parameters.....	77
Table B.4 : k-Nearest Neighbors algorithm parameter.....	77
Table B.5 : Support Vector Machines algorithm parameters.....	78
Table B.6 : Extreme Gradient Boosting algorithm parameters.	78

LIST OF FIGURES

	<u>Page</u>
Figure 2.1 : Squirrel cage induction motor structure, courtesy of WAT Motor Co.	6
Figure 2.2 : Maintenance types, adapted from [2].....	8
Figure 2.3 : The P-F curve shows the point where the fault started, became observable and the fault occurred, adapted from [3].	9
Figure 2.4 : Star-connected stator winding faults, adapted from [4].....	13
Figure 2.5 : Flowchart of Power Spectral Density estimation via Welch's method [5].....	22
Figure 2.6 : Algorithm of Power Spectral Density estimation via Welch's method [6].....	23
Figure 2.7 : An example of Confusion Matrix.	29
Figure 3.1 : Experiment setup, courtesy of WAT Motor Co.....	33
Figure 3.2 : Typical induction motor label, courtesy of WAT Motor Co.	34
Figure 3.3 : Schematic of test system.....	35
Figure 3.4 : An example of stator current signals of the healthy and bearing-fault motor at rated load.....	36
Figure 3.5 : An example of stator current signals of the healthy and bearing-fault motor at 75% of the rated load.....	37
Figure 3.6 : An example of stator current signals of healthy and stator inter-turn-fault motor at 75% of the rated load.....	38
Figure 3.7 : An example of stator current signals of healthy and stator inter-turn-fault motor at 75% of the rated load.....	39
Figure 3.8 : An example of stator current signals of healthy and broken rotor bar-fault motor at 75% of the rated load.....	40
Figure 3.9 : An example of stator current signals of healthy and broken rotor bar-fault motor at 75% of the rated load.....	41
Figure 3.10 : Welch's PSD estimations of the healthy and bearing-fault motor at rated load.....	43
Figure 3.11 : Welch's PSD estimations of the healthy and bearing-fault motor at 75% of the rated load.	44
Figure 3.12 : Welch's PSD estimations of the healthy and stator inter-turn-fault motor at 75% of the rated load.....	45
Figure 3.13 : Welch's PSD estimations of the healthy and stator inter-turn-fault motor at 75% of the rated load.....	46
Figure 3.14 : Welch's PSD estimations of the healthy and broken rotor bar-fault motor at 75% of the rated load.....	47
Figure 3.15 : Welch's PSD estimations of the healthy and broken rotor bar-fault motor at 75% of the rated load.....	48
Figure 4.1 : Flowcharts of the methods presented in this thesis.....	50
Figure 4.2 : Diagram of time-domain statistical analysis method.....	51

Figure 4.3	: t-SNE plot of time-domain statistical features.....	51
Figure 4.4	: Diagram of frequency domain statistical analysis method.	52
Figure 4.5	: t-SNE plot of frequency domain statistical features.	52
Figure 4.6	: Diagram of statistical analysis of characteristic frequencies method.	53
Figure 4.7	: t-SNE plot of characteristic frequencies statistics.	54
Figure 4.8	: Diagram of deep learning analysis.....	56
Figure 4.9	: Sliding window data augmentation.....	56
Figure 4.10a:	Training performance of 1D-CNN model.	59
Figure 4.10b:	Training performance of LSTM model.	59
Figure C.1	: 1D-CNN model diagram.....	80
Figure C.2	: LSTM model diagram.....	81

DATA-DRIVEN CONDITION MONITORING AND FAULT DIAGNOSIS OF VFD-FED INDUCTION MOTORS

SUMMARY

The impact of Industry 4.0 awakens the need for equipment and system-level condition monitoring, especially for asynchronous motors, which are the backbone of the industry, in parallel with the increasing search for efficiency in industrial operations and the importance of data. Although the number of asynchronous motors fed directly from the line still dominates, the number of motors driven by VFD is also increasing with the demand for efficiency and the regulations applied.

Within the scope of this thesis, the most common faults in asynchronous motors, bearing, stator winding short circuit, and broken rotor bar faults were created in the laboratory environment of WAT Motor Company facilities, condition monitoring and fault diagnosis studies were carried out. Studies on line-start asynchronous motors are widely applied in the literature. However, in this study, tests were carried out in two different loadings at different speeds between 30 Hz and 50 Hz at 75% of the nominal and nominal load of the motor fed with the variable frequency drive.

With the obtained data, different methods and approaches are examined for motor condition monitoring and fault diagnosis. In the first method, the statistical features of the current signal are extracted for time-domain analysis. In the study conducted in the frequency domain, the statistical features of the frequency spectrum amplitudes obtained by estimating the Power Spectral Density with the Welch method were obtained in the same way. Finally, the amplitudes corresponding to the characteristic frequencies of the fault types in the frequency spectrum were calculated and the statistical properties of these amplitudes were extracted. With the features obtained by these three methods, machine learning classifiers were trained and motor fault diagnosis was performed. According to the results obtained, it is possible to achieve good results in fault diagnosis with statistical approaches. On the other hand, considering the industrial application conditions, electric motors are heavily exposed to disruptive effects. PSD estimation with the Welch method gives very robust results against disturbance effects due to its nature. The amplitude-based statistical approach obtained from the PSD frequency spectrum with error characteristic formulas overperformed the other two methods with high accuracy and precision in all metrics.

Increasing deep learning studies and applications in recent years find a place for themselves in the diagnosis of motor failures. As an alternative to feature extraction and signal processing problems that require time and expertise in machine learning methods, deep learning methods can offer an end-to-end solution. As an alternative to the data engineering and machine learning methods mentioned in the thesis, two

deep learning methods, convolutional and recurrent neural networks structured for fault diagnosis. Deep learning methods offer high performance without any preprocessing, but as a disadvantage, they require high processing power and a large dataset to train the model. Although deep learning methods need large data sets in the training phase, they can work with less dimensional data than machine learning methods, as shown in the thesis.

**DEĞİŞKEN FREKANSLI SÜRÜCÜ
İLE BESLENEN ASENKRON MOTORLarda
VERİ ODAKLI DURUM İZLEME VE ARIZA TANILAMA**

ÖZET

Elektrik motorları, sadece endüstriyel uygulamalarda değil, aynı zamanda konut, tarım ve ulaşım amaçlı olarak da elektrik gücünü mekanik güçe dönüştüren sistemlerde yaygın olarak kullanılmaktadır. Sürdükleri sistemlerle birlikte ele alındığında, elektrik motorları tüm elektrik tüketiminin %40'ından fazlasını ve bir sonraki en büyük tüketici olan aydınlatmanın neredeyse iki katı kadarını kullanır. Sadece endüstriyel kullanım düşünüldüğünde, elektrik motorları toplam elektrik tüketiminin %70'ine yakınına oluşturmaktadır.

Endüstriyel uygulamalarda birçok farklı motor tipi mevcuttur ancak alternatif akım (AC) asenkron motorlar basit, güvenilir ve sağlam tasarımları nedeniyle en çok tercih edilen tiptir. Malzeme taşıma, malzeme işleme, pompalama, havalandırma ve basınçlı hava üretimi gibi temel elektro-mekanik sistemleri çalıştırın AC asenkron motorların nispeten kayip maliyeti, düşük bakım, yüksek güvenilirlik ve uzun kullanım ömrü en avantajlı özellikleridir. Özellikle HVAC (Isıtma, havalandırma ve iklimlendirme) sektörü, endüstriyel elektrik tüketiminde en büyük paya sahip oldukları ve oldukça yüksek tasarruf potansiyellerine sahip oldukları için özel önem gerektirmektedir.

Son yıllarda küresel ısınma konusunda artan farkındalık, elektrik motorlu sistemler de dahil olmak üzere daha verimli sistemler talep ediyor. Avrupa Parlamentosu ve Avrupa Konseyi gibi otoriteler, yüksek verimli premium motorların ve değişken frekanslı sürücülerin (VFD) kullanımını teşvik ederek verimliliği artırmak için yeni gereksinimler uygulamaktadır. VFD'ler, motorun çıkış torkunu ve hızını mekanik sistem yüklerine uyacak şekilde düzenler ve pompalar, fanlar ve kompresörler gibi yüksek düzeyde doğrusal olmayan giriş gücüne, çıkış torku ve hızına sahip değişken mekanik gücün gerekli olduğu yerlerde önemli enerji verimliliği sağlar.

20 yıllık kullanım süresi göz önüne alındığında, bir elektrik motorunun güç tüketimi toplam sahip olma maliyetinin %90'ını oluştururken, bunu %5 ile arıza süresi maliyeti ve %4 ile bakım maliyeti izlemektedir. İlk satın alma fiyatı ise toplam maliyetin sadece %1'ini temsil ettiği düşünüldüğünde, motorun çalışması sırasında alınan önlemlerle tasarruf sağlanabileceği sonucuna varılabilir. Bu noktada ise Endüstri 4.0, otomasyon ve verimlilik yoluyla endüstriyel operasyonları şekillendiriyor. Durum izleme, çalıştığı süre boyunca tesisin ve/veya ekipmanın durumunu değerlendirerek Endüstri 4.0'ın temel yapıtaşlarından birini oluşturmaktadır.

Bakım, tüm kullanım ömrü boyunca tasarlanan işlevlerini sürdürmek için ekipmanı korumak veya eski haline getirmek için yapılan eylemler olarak tanımlanabilir. Geleneksel bakım, çalışabilirliği sağlamak için periyodik sağlık kontrollerine dayanır,

ancak araştırmalar, bakım zamanında ve doğru bir şekilde yapılsa bile arızaların büyük çoğunluğunun çalışma durumunda ortaya çıktığını göstermektedir. Durum izleme ve arıza teşhisini, bu tür durumları önlemek için bakım planlamasına yardımcı olurken, istenmeyen arıza sürelerini ve mali kayıpları da önler. Ayrıca durum izleme, uzun vadeden daha güvenilir bir sistem sağlayan ekipman veya tesisin trend analizi yoluyla daha iyi anlamak için bir veritabanı oluşturma fırsatına da sahiptir.

Durum izleme, arıza tespiti için bir teşhis aracı ve bakım planlamasının temellerinden biri olarak motora sürekli veya periyodik olarak uygulanır. İzlenen parametrelerdeki ani veya beklenmedik değişiklikler motorun durumu hakkında önemli bilgiler sağlar. Rulmanlardan elektrik motorlarına ve pompalara kadar çeşitli ekipmanların sağlığına ilişkin bilgileri değerlendirmek için kullanılabilecek titreşim, sıcaklık ve akım izleme gibi birçok durum izleme yöntemi mevcuttur. Akım izleme, asenkron motor çalışmasını kontrol etmek için kolayca ölçülebildiği için kendisini diğer yöntemlerden ayırmaktadır.

Besleme akımı sinyalleri aracılığıyla durum izleme, yalnızca motorun kendisi için değil aynı zamanda motorun tahrik ettiği mekanik sistem hakkında da faydalı bilgiler sağlar. Bakım stratejisinin önemli bir yönü, motorun tahrik ettiği mekanik bir sistemin dahil edilmesidir. Özellikle motorun çalışmasını kontrol etmek için akımın zaten algılandığı VFD beslemeli sistemlerde, ek sensör ihtiyacı olmadan hem elektriksel hem de mekanik arızalar teşhis edilebilir. Akım sinyallerini kullanarak arıza teşhisini için birçok çalışma yapılmışmasına rağmen, VFD beslemeli motorlarla ilgili çalışmalar sınırlıdır. VFD’lerde kullanılan PWM sinyallerinin, motor akım sinyalindeki bir arızanın özelliklerini maskeleyerek, teşhisini zorlaştıracabileceğine dikkat edilmelidir. Bu çalışmada, VFD beslemeli üç fazlı asenkron motorun tek fazlı stator besleme akımı üzerinden farklı yük ve frekans senaryolarında elektriksel ve mekanik arıza tespiti üzerinde durulmuştur.

Bu çalışmada öncelikle asenkron motorların endüstriyel alandaki yeri ve önemi ile genel çalışma prensibinden bahsedilmiştir. Literatür taraması kapsamında işletme ömürleri boyunca maruz kaldıkları stresler nedeniyle oluşan arızalar açıklanmıştır. Arızaların olumsuz sonuçlarından en az hasarla kaçınmak için etkin bakım yöntemlerinden bahsedilmiş ve verimli bakımın omurgası olarak durum izleme ve hata tanıma yöntemleri anlatılmıştır. Çalışma kapsamında asenkron motor arızalarının en sık görülen türleri olan rulman, stator sargı kısa devre ve kırık rotor çubuk arızaları WAT Motor tesislerinde laboratuvar koşullarında oluşturulmuş ve etkileri incelenmiştir. Literatürde besleme hattından doğrudan beslenen ve nominal hızda çalışan asenkron motorlar ile ilgili çalışmalar yaygındır. Bu çalışmada, değişken frekanslı sürücü ile beslenen motorun nominal ve nominal yükünün %75’inde 30 Hz ile 50 Hz arasında farklı hızlarda iki farklı yüklemede testler gerçekleştirilmiştir.

Elde edilen veriler ile motor durum izleme ve arıza tespiti için farklı yöntem ve yaklaşımlar incelenmiştir. Bunlardan ilkinde, zaman domeninde yapılan analizde akım sinyalinin istatistiksel özellikleri çıkarılmıştır. Frekans alanında yapılan çalışmada ise, Welch yöntemi ile Güç Spektral Yoğunluğu tahmin edilerek elde edilen genliklerin istatistiksel özellikleri aynı şekilde elde edilmiştir. Son olarak frekans spektrumundaki arıza tiplerinin karakteristik frekanslarına karşılık gelen genlikler hesaplanmış ve bu genliklerin istatistiksel özellikleri çıkarılmıştır. Bu üç yöntemle elde edilen özellikler ile makine öğrenmesi sınıflandırıcıları eğitilmiş ve motor arıza teşhisini gerçekleştirilmiştir. Elde edilen bulgulara göre istatistiksel yaklaşımalarla arıza teşhisini

konusunda iyi sonuçlar almak mümkündür. Endüstriyel uygulama koşulları göz önüne alındığında elektrik motorları, bozucu etkilere yoğun bir şekilde maruz kalmaktadır. Welch yöntemi ile PSD tahmini ise, yapısı gereği bozucu etkilere karşı oldukça sağlam sonuçlar vermektedir. Hata karakteristik formülleri ile PSD frekans spektrumundan elde edilen genlik tabanlı istatistiksel yaklaşım, tüm metriklerde yüksek doğruluk ve kesinlik ile diğer iki yöntemi geride bırakmıştır.

Son yıllarda artan derin öğrenme çalışmaları ve uygulamaları, motor arızalarının teşhisinde de kendisine yer bulmaktadır. Makine öğrenmesi yöntemlerinde zaman ve uzmanlık bilgisi gerektiren özellik çıkarma ve sinyal işleme problemlerine alternatif olarak derin öğrenme yöntemleri uçtan uca bir çözüm sunabilmektedir. Tez kapsamında bahsedilen veri mühendisliği ve makine öğrenmesi yöntemlerine alternatif olarak evrişimli ve yinelemeli sinir ağları olmak üzere iki derin öğrenme yöntemi ile hata teşhisini yapmıştır. Derin öğrenme yöntemleri, herhangi bir ön işleme gerekmeden yüksek performans gösterir, ancak bir dezavantaj olarak, modeli eğitmek için yüksek işlem gücüne ve büyük bir veri kümesine ihtiyaç duyarlar. Derin öğrenme yöntemleri, eğitimleri sırasında büyük veri kümelerine ihtiyaç duysa da tez kapsamında gösterildiği gibi makine öğrenmesi yöntemlerine göre daha az boyutlu verilerle çalışabilirler. Bu açıdan VFD'ye entegre edilecek eğitilmiş bir derin öğrenme algoritması, ileri beslemeli yapısı sayesinde daha az kaynak ihtiyacı ile çalışabilir.

Sonuç olarak Endüstri 4.0 kapsamında bazı kavramlar daha da önemli hale geliyor. Bunlardan ilk ikisi verilere erişimi kolaylaşması ve veri işleme gücünün artması iken, üçüncüsü verimlilik olarak sayılabilir. Asenkron motorlarda bu üç kavram VFD'ler ile karşılanabilir. VFD ile beslenen motorlarda sistemin verimi artarken, motor kontrolü için gerekli olan akım sinyali herhangi bir ek masraf olmadan durum izleme çalışmaları için kullanılabilir. Toplanan verileri Nesnelerin İnterneti altyapıları ile veri merkezlerine aktararak veri erişimini kolaylaştırma potansiyeline de sahiptir.

1. INTRODUCTION

1.1 Overview

Electric motors are extensively employed in a system that converts electrical power into mechanical power in not only industrial applications but also residential, agricultural and transportation purposes. Taken together with systems they drive, electric motors use more than 40% of all electricity consumption and almost twice as much as the next largest user lighting [7]. Considering only industrial usage, electric motors dominate and account for close to 70% of the total electricity consumption [7, 8].

There are many different motor types available in industrial facility operations, but asynchronous alternating current (AC) induction motors are the most preferred type because of their simple, reliable and rugged design. Relatively low cost, low maintenance, high reliability and long lifespan are the most advantageous features of AC induction motors which drive core electro-mechanical systems such as material handling, material processing, pumping, ventilation and compressed air generation [9]. Especially HVAC (Heating, ventilation and air conditioning) sector requires special attention as they have the largest share of industrial electrical consumption and reasonably high saving potentials [9].

In recent years raised awareness about global warming demands more efficient systems including electric motor-driven systems. Policymakers such as the European Parliament and the European Council implementing new requirements to increase efficiency by encouraging the usage of high-efficiency premium motors and variable frequency drives (VFD) [8, 10].

VFDs regulate the motor's output torque and speed to match the mechanical system loads and enable significant energy efficiency where variable mechanical power is needed that have highly non-linear input power, output torque and speed such as pumps, fans and compressors. Previously Direct Current (DC) motors have

been dominant for variable motor speed control, yet developments in semiconductor technology became the driving force behind the prevalence usage of VFDs with AC motors [11]. Motor speed control is advantageous in terms of lower system energy costs, increased system reliability and less maintenance.

Considering 20-year in service, the power consumption of an electric motor depicts 90% of the total cost of ownership and followed by downtime costs as 5% and rebuild costs as 4% [7]. The initial purchase price represents only 1% of the total cost and it can be concluded that savings can be achieved by actions taken during operation of motor [7].

Industry 4.0 is shaping industrial operations through automation and efficiency. Condition monitoring paves the way to Industry 4.0 through evaluating the state of the plant and/or equipment throughout its service life [2]. Maintenance can be defined as actions to retain or restore equipment in order to maintain its designed functions within the entire lifespan [2]. Traditional maintenance relies on periodically health checks to provide operability, but researches show that even if maintenance is done on time and correctly the vast majority of failures arises during operation state [12]. Condition monitoring and diagnostics can help to schedule maintenance to prevent such situations whilst avoiding unintended downtime and financial losses. Also, condition monitoring has the opportunity to build a database to understand better via trend analysis of the equipment or plant that leads more reliable system in the long run.

There are many condition monitoring methods available such as vibration, temperature, and current monitoring that can be used to assess insights into the health of equipment varying from bearings to electric motors and pumps. Current monitoring distinguishes itself from other methods since it is readily measured to control induction motor operation. VFDs are presenting a great potential not only to control the motor operation but also to be utilised as a connection to the Internet of Things structure to serve Industry 4.0.

1.2 Objectives of Research

This study aims to diagnose and identify mechanical and electrical faults of VFD-fed induction motors under various loads and speeds via monitoring only motor current. As an outcome of this research comparative results among time-domain versus frequency-domain analysis and machine learning algorithms versus deep learning algorithms are presented.

The achievement of this study was facilitated by the following specific objectives:

- Analyse motor faults under VFD controlled motor current
- Investigate effects of various loads and speeds
- Build different feature engineering methods
- Benchmark Machine and Deep Learning algorithms

1.3 Organization of Thesis

Thesis organised in five chapters to achieve aforementioned objectives;

- Chapter-2 provides an in-depth background to condition monitoring and fault diagnosis of AC induction motors including general information about induction motors, fault types, condition monitoring and signal processing techniques followed by fault diagnosis methods and performance metrics.
- Chapter-3 presents the experimental testing system and used methodology.
- Chapter-4 discusses the diagnostics of faults via two different approaches: machine learning and deep learning-based condition monitoring.
- Chapter-5 remarks obtained results with different approaches and concludes with future recommendations.

2. CONDITION MONITORING OF INDUCTION MOTORS: BACKGROUND

2.1 Introduction of Induction Motors

2.1.1 Principle of operation

Electric motors are divided into two classes depending on their power supply type: direct current (DC) or alternating current (AC). The latter can be broken into two classes as synchronous or induction according to their operating speed. Induction motors, which operates slightly lower than synchronous speed, are also sub-divided as wounded and squirrel-cage motors. In this study, squirrel-cage induction motors have been investigated by means of induction motors, since the squirrel-cage type is predominantly used in industrial applications.

Induction motors run at a speed slightly slower than the synchronous speed at the point where motor torque and load torque are equal [13]. The difference between the actual rotor speed and synchronous speed is called slip [11].

$$\text{Synchronous Speed} = \frac{120 \cdot \text{Frequency (Hz)}}{\text{number of poles}} \quad (2.1)$$

$$\text{Slip} = \frac{\text{Synchronous Speed} - \text{Rotor's Mechanical Speed}}{\text{Synchronous Speed}} \quad (2.2)$$

In Principle, induction motors transfer electrical energy into mechanical energy by interlinking two electrical components: stator as stationary part and rotor as rotational part. Electrical energy is transmitted from stator to rotor via electromagnetic induction, then a mechanical component bearing guides rotor to provide mechanical power [4,14].

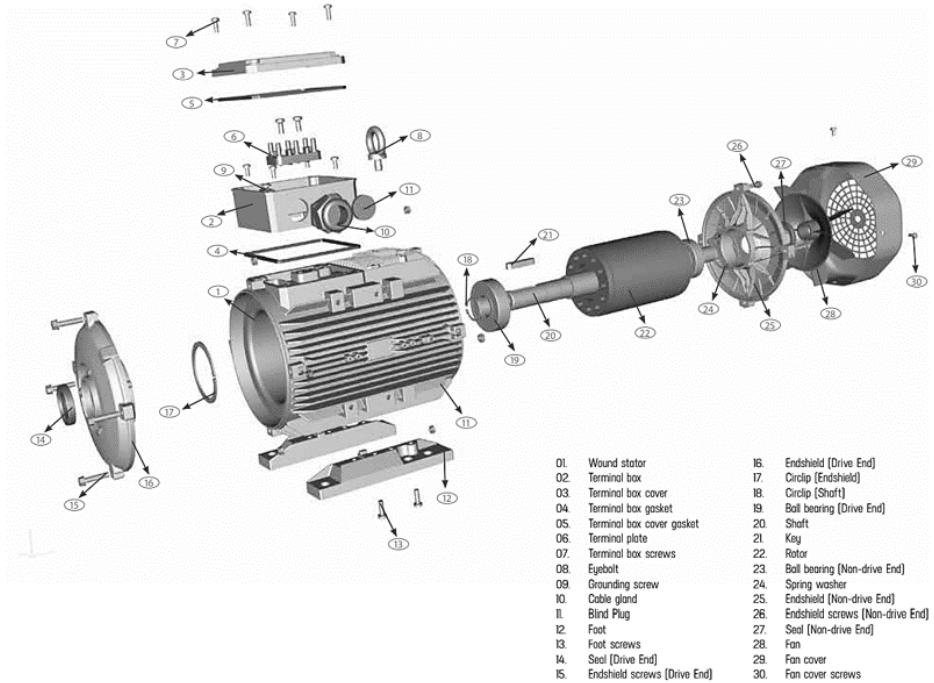


Figure 2.1 : Squirrel cage induction motor structure, courtesy of WAT Motor Co.

2.1.2 VFD-fed induction motors

A variable frequency drive, also named as the inverter, adjustable-frequency drive (AFD) or variable-speed drive (VSD), fed motor system controls the rotation speed of the induction motor by regulating the supply frequency and voltage of the motor. The main difference between line-start and VFD-fed induction motors is that while in line-start mode supply voltage is the only controllable parameter, on the other hand, VFD-fed has the ability to control torque and speed easily [15].

From a historical point of view, DC motors have been utilised in speed control applications. However, as a result of advances in power semiconductor technology used in inverters, the performance of AC motors in terms of precision, response, and speed range began to exceed that of DC motors [10, 11]. As a driving force behind the induction motor control dominance today, VFDs generally have the following control strategies regarding speed and torque regulation [16, 17]:

- Voltage per Frequency Control (V/f)
- Field Oriented Control (FOC)
- Direct Torque Control (DTC)

The common idea behind these methods is based on controlling the torque and flux references applied to the motor separately, as in DC motor control [15]. In the scope of this thesis, only the V/f control strategy is emphasized due to the widespread adoption of the control method in pump, compressor and fan applications.

V/f control can be employed in both open-loop and closed-loop modes. Open-loop V/f control, which is by far the most popular control due to its simplicity, as the name implies, creates a constant air-gap flux by detaining the ratio between the voltage and frequency applied to the induction motor constant, and as a result, it provides the opportunity to work at operating frequencies from zero to nominal frequency [18].

VFDs come with benefits such that energy savings, reliability and product quality, yet in concern of fault diagnosis they introduce a number of factors, which will be discussed later on, that increase the complexity.

2.1.3 Need for condition monitoring

Condition monitoring is defined as measuring activities concerning characteristics and parameters of physical equipment at predetermined intervals either manually or automatically [2]. Leveraging rapid technological advancements in data storage, data process and network structure, condition monitoring became one of the driving force behind the industry 4.0 paradigm. The key goal behind this paradigm is to acquisition, transmission and analysis of data in order to predict the future behaviour of machinery, or plant on a larger scale, to boost efficiency and reliability [19, 20].

Researchers from both academia and industry have devoted significant attention to condition monitoring of induction motors over decades. Even though induction motors are renowned for robustness, environmental, electrical and mechanical effects may lead induction motors to failure. As a result, industrial processes are subjected to potential losses in a manner of time and capital, so the desire to minimize or even prevent these losses emerges the need for condition monitoring.

2.1.4 Maintenance strategies

Maintenance can be defined as the combination of all technical and managerial actions taken to maintain or restore an item throughout its life cycle in a condition where it can fulfil its designed function [2]. A motor maintenance program should effectively address reliability, cost, and scheduling issues, as well as the causes of the most common motor failures. Essentially, there are two types of maintenance strategies: corrective and preventive.

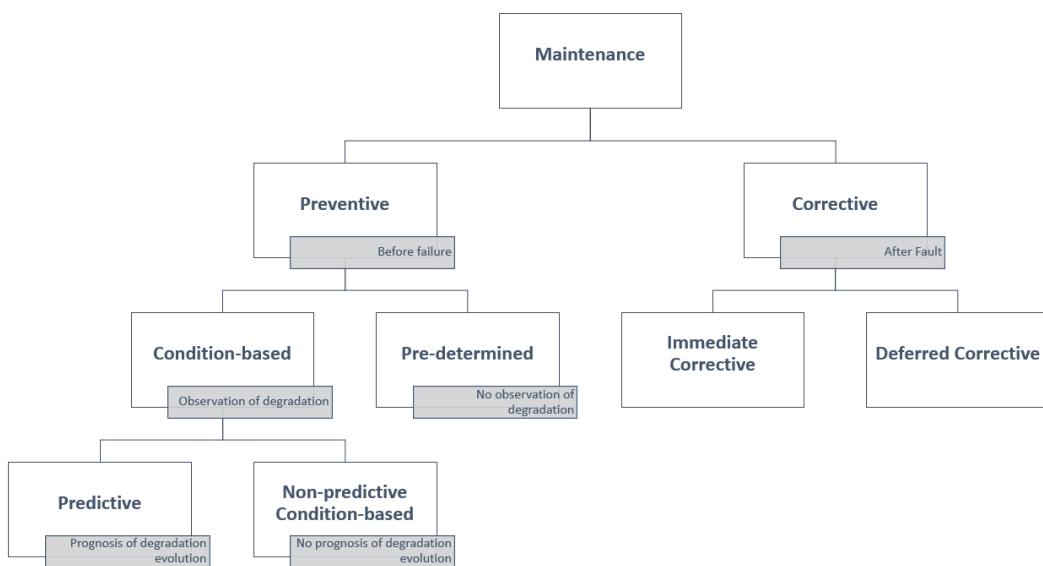


Figure 2.2 : Maintenance types, adapted from [2].

Corrective maintenance is a form of maintenance conducted after the induction motor failure to detect the fault and restore it to operational condition [2]. The main purpose of this type of maintenance is to get the equipment up and running as soon as possible by repairing or replacing the defective equipment. However, corrective maintenance as a failure-driven method contains a high-risk potential as faults may occur at unexpected times, can disrupt the operation. Since this type of maintenance approach does not take into account the damages that may occur, it may be suitable for equipment that is not critical to the business that does not pose a safety risk.

Preventive maintenance, on the other hand, aims to detect faults at an early step and correct them before they create risk to operation [2]. Preventive maintenance employed to increase efficiency and reliability by taking into account the probability of failure or the ageing of the equipment, at certain intervals or according to

pre-planned scheduling. Although this approach is beneficial in cases where the wear-out characteristics are evident, it has disadvantages, especially in terms of not being able to use equipment lifespan efficiently and increasing the maintenance cost compared to the corrective maintenance approach [21].

Predictive maintenance is a condition-based approach to maintenance that is used to evaluate the parameters and characteristics of the equipment or to make predictions based on repeated analysis [2]. Compared to preventive maintenance, predictive maintenance maximizes equipment service life whilst minimizing unnecessary maintenance. In 99% of machine failures, it is possible to observe indications that malfunctions will occur, in other words, the necessary measures can be taken before 99% of the faults occur by continuously monitoring the machine [21].

Within the predictive maintenance approach, decision-making can be divided into two: diagnosis, which is the analysis of the current situation, and prognosis, which is the assessment of conditions measured over time [3]. A P-F curve can be used to better understand diagnostic and prognostic monitoring systems.

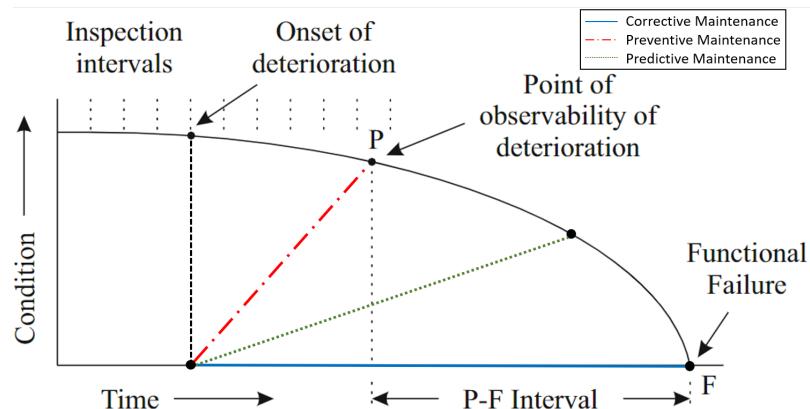


Figure 2.3 : The P-F curve shows the point where the fault started, became observable and the fault occurred, adapted from [3].

The downside of predictive maintenance is that it requires additional equipment and relatively high investment costs. But the advantage of VFDs also comes out here. As they currently monitor motor parameters in control applications, they have a high potential for predictive maintenance applications without the need for additional sensors and investments.

2.2 Induction Motor Fault Types

From a mechanical perspective, induction machines basically consist of three components: stator, rotor and bearing. Electrical, mechanical, and environmental disturbances constantly affect asynchronous motor components and cause most malfunctions [22]. Table 2.1 exhibits various surveys that studied and categorized the most common failures [1, 12, 23–25].

Table 2.1 : Distribution of induction motor faults by component (%).

Component	IEEE	EPRI	Thorsen-Dalva	Bonnett-Yung
Bearing	44	41	51	69
Stator	26	37	16	21
Rotor	8	10	5	7
Other	22	12	28	3

Considering the actual distribution of the faults, in total 80% of the faulty motors have only one fault, while this rate reaches 90% especially in low voltage supply motors [1].

Table 2.2 : Actual distribution of multiple failures (%) [1].

Exact Failure	Motor Supply Voltage			Total
	Low	Middle	High	
1 Fault	91.1	79.6	77.1	79.9
2 Faults	8.0	11.8	13.1	11.9
3 Faults	0.9	3.3	5.5	4.0
4 Faults	0.0	2.2	0.8	1.3
> 4 Faults	0.0	3.0	3.5	2.9

As presented in Table 2.1, most of the faults are associated with bearings followed by stator related faults. It also should be noted that these surveys do not include the effects of power electronics. A motor controlled by VFD is subjected to short and high voltage pulses called PWM (Pulse Width Modulation), which are sent at a very high frequency, which can have a detrimental effect on the wire insulation and cause a burn on the stator [13]. Although this problem can be solved with high-quality insulation, PWM signals also create non-continuous electrical discharges on the bearings, causing wear which reduces bearing lifespan [26]. Therefore, it would not be wrong to conclude that bearing and stator failures will also have a high rate in VFD-fed induction motors.

2.2.1 Bearing related faults

In all kinds of electrical machines, the mechanical element positioned between the frame that initiates the movement and the rotating axis shaft is called a bearing. These mechanical elements, which help the rotational movement of the electric motor, are exposed to many internal and external destructive effects during their operation and failures arise as a result. Major sources of bearing failures are given below [22,27–31]:

Mechanical stresses: Fatigue, which mostly begins on the surface, turns into small-sized material ruptures at the beginning and later dimensional surface indentations and protrusions. Loose motor connection, misalignment where the motor shaft and load shaft are connected without aligning on the same axis, angular misalignment where the motor shaft and load shaft axes are connected at a certain angle, and unbalanced load connection, which is an unbalance condition where the centre of gravity of the load connected to the motor shaft is not on the rotation axis are other mechanical disturbances on the bearing.

Environmental stresses: Corrosion occurs on the bearing surfaces used in high humidity working environments. Especially the moisture absorbed in the bearing oil initiates this process and the rust that occurs due to corrosion causes deterioration that turns into indentation and protrusion on the surface of the bearing element, and cracks in the later stages.

Thermal stresses: Insufficient lubrication generally causes problems with bearing components. Normally, there is a layer of oil in the bearing that prevents direct contact between the rotating elements so that their surfaces do not wear out quickly. In case of insufficient lubrication, excessive wear and subsequent material deterioration occur as a result of increased friction due to direct contact between metal surfaces.

Electrical stresses: the electrical discharge current effect occurs with a fault current flowing through the bearings from the motor frame to the ground in motors that do not have a suitable ground connection. Asymmetry of stator windings, permanent magnetism effect developing in the motor over time, electrostatic charge accumulation in the motor frame and application of voltage to the motor shaft from the outside, or common end voltages generated due to the high switching frequency of semiconductor power electronics (VFDs using PWM) are the factors that cause this malfunction. The

irregular current will cause wear and tear on the bearing metal surface, and as a result, the degree of material rupture and surface deterioration increases.

Vibration in the motor causes the rotor to rotate irregularly or axially unbalanced in the motor air gap. Any axial misalignment that occurs in the motor air gap adversely affects the air gap flux density and causes the formation of harmonic components [15, 30, 31]. Consequently, this can induce harmonic components in the current drawn by the motor with frequencies given by formula [30]:

$$f_{bng} = f_e \pm m \cdot f_v \quad (2.3)$$

where,

f_e is the electrical supply frequency;

f_v is the rotational speed frequency of the rotor;

m is the harmonic number 1, 2, 3,;

f_{bng} is the current component frequency due to air gap changes.

2.2.2 Stator related faults

As researches have shown, stator faults occupy an important place among asynchronous motor faults after bearing [1, 12, 23–25]. Mechanical, electrical, thermal and environmental factors cause malfunctions in the stator windings, as well as their laminations [4, 32]. Winding faults, as the most common stator faults, are winding short-circuit faults that are mostly the result of the aforementioned effects of the winding insulation. Types of winding faults are as follows [4, 32, 33]:

- Short-circuit between two turns in the same phase (turn-turn failure),
- Short-circuit between two coils side by side in the same phase (coil-coil failure),
- Short-circuit between the turns of two phases (phase-phase failure),
- Short-circuit consisting of all three-phase turns,
- Short-circuit between the conductor of the winding and the stator core (phase-ground short circuit),
- Open-circuit fault when winding breaks.

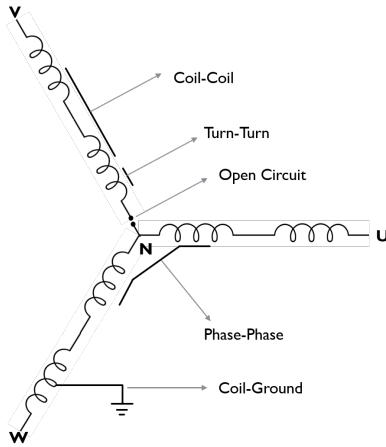


Figure 2.4 : Star-connected stator winding faults, adapted from [4].

The factors that cause the motor winding insulation to deteriorate are explained below [4, 15, 32, 34]:

Mechanical stresses: While the motor is running, the rotor may rub or hit the inner surface of the stator due to motor shaft deterioration, bearing failures and misalignment. This force creates a turn-to-turn or a phase-to-earth short-circuit, causing the stator coil and the stator winding insulation to break down. On the other hand, winding breakage may occur due to vibration during operation and therefore the motor produces the open-circuit fault.

Environmental stresses: The environment in which the motor is running can be very hot, cold or humid. On the other hand, substances in the external environment can contaminate the windings, causing the heat dissipation to deteriorate and the insulation to be damaged. In addition, the airflow can be blocked and cannot absorb the air required for cooling. Therefore, it causes the motor windings to heat and consequently the insulation to deteriorate.

Thermal stresses: Thermal effects occur as a result of overloading or a motor failure. With motor overload, the motor temperature rises above the limit value of the insulation class and the insulation deteriorates. At this point, every 3.5% unbalance in the motor supply voltage increases the temperature of the motor by 10°C. In addition, every 10°C temperature increment above the limit temperature value of the insulation halves the life of the insulation.

Electrical stresses: The main reason for this is sudden changes in supply voltage. Transients during commissioning and decommissioning and voltage fluctuations

frequently occur, especially in asynchronous motors powered by variable frequency drives. Winding insulations deteriorate due to these voltage variations.

Under the inter-turn short-circuit condition, a significant deviation in rotor slot harmonics components, called as principle slot harmonics (PSH), occurs and can be obtained by given formula [35];

$$f_{st} = f_e \cdot \left[n \cdot \frac{(1-s)}{p} \pm m \right] \quad (2.4)$$

where,

f_e is the electrical supply frequency ;

p is the number of pole pairs of the motor ;

$n = 1, 2, 3 \dots (2p - 1)$;

s is the slip ;

m is the harmonic number $1, 2, 3, \dots$;

f_{st} is the principle slot harmonic frequencies.

2.2.3 Rotor related faults

There are several reasons why rotor bar faults can occur in an induction motor. In caged motors, breaking one cage bar does not significantly change the operating behaviour of the machine. However, due to the fracture that occurs, the current distribution, air gap flux, force balance and temperature distribution in the rotor deteriorate, and heating and strains increase [36]. If the rotor continues to run in this state, damage can also spread to the sidebars, causing multiple bars of the rotor to break. In this respect, it is very important to diagnose the condition when a rotor bar is broken.

The main reasons of rotor broken bar of an induction motor can be listed as follows [32, 34, 36, 37];

Mechanical stresses: In cases that cause structural asymmetry such as rotor misalignment or bearing failure, the resultant of the normal direction forces in the air gap is not equal to zero and the force acting on the bars increases. In addition, dynamic effects such as impact forces due to sudden load change, centrifugal forces due to excessive acceleration also cause failure.

Environmental stresses: Dusty, wet and/or oily environment in which the electric motor operates negatively affects the engine and increases the possibility of malfunction.

Thermal stresses: Thermal stresses may occur during take-off and/or operation. The temperature limit values of the motor and rotor are different. In terms of the safe operation of the motor, the rotor temperature at start-up and the stator temperature during operation are decisive. Thermal stresses are generally caused by frequent starting, locking of the motor shaft, bearing failure, insufficient cooling, skin effect and current accumulation. It takes the form of partial warming in machines fed by power electronics.

Electrical stresses: The flux created by the current flowing through the rotor bars creates an electrodynamic force ($F \propto I^2$) acting from the rotor surface towards the shaft in quadratic proportion to the current. The bar vibrates at $2 \cdot s \cdot f_e$ and $4 \cdot s \cdot f_e$ frequencies and can therefore cause breakage in the rotor bars. In addition, since the rotor current at motor start-up is very high, the rotor bar is again exposed to high stresses.

Cracked or broken bar in the rotor cage results a series of sideband frequencies in the stator current given by [4];

$$f_{brb} = f_e \cdot [1 \pm 2 \cdot m \cdot s] \quad (2.5)$$

where,

f_e is the electrical supply frequency ;

s is the slip ;

m is the harmonic number 1, 2, 3;

f_{brb} is the broken rotor bar sideband frequencies.

2.3 Condition Monitoring Techniques

Condition monitoring is applied to the motor continuously or periodically, as a diagnostic tool for fault detection and as one of the fundamentals of maintenance planning. Sudden or unexpected changes in monitored parameters yield important information about the condition of the motor. Although the parameters to be monitored

vary depending on the end-user, temperature, vibration and current magnitudes are widely used in the industry [38].

2.3.1 Temperature monitoring

One of the parameters that can be followed in order for electric motors to work safely and without failure for a long time is the motor temperature. By installing sensors such as Resistance Temperature Detectors, thermistors, thermocouples and thermostats, electrical or mechanical faults can be detected [38]. These sensors are usually placed in the stator windings, bearings and frame [39]. In addition, temperature monitoring can be performed by parameter estimation over the stator supply current without using any temperature sensor [40].

As mentioned before, thermal stresses resulting from effects such as overloading and bearing lubrication problems may damage various components of the motor. While bearing temperatures maintain useful information about possible friction problems, the coolant bulk outlet temperature is frequently monitored, especially when the machine is forced above its nominal values, and winding temperature monitoring is also useful in the event of overheating due to overload [41]. Continuously monitoring temperature will give an indication of potential failures to avoid catastrophic incidents.

2.3.2 Vibration monitoring

Due to their working principles, rotating gears, electric fields and shafts periodically generate vibrations [42]. Since the produced vibration signals contain information about the condition of the machine and can be followed without interfering with the operation of the motor, it is mostly preferred in condition monitoring studies. Vibration monitoring has the ability to track sudden changes in the machine condition that enables monitor the condition of the equipment continuously or intermittently.

Vibration can be measured in units of displacement, velocity, and acceleration. The displacement type is generally used in the measurement of rotor vibration, while the velocity type is used in motor housing vibration measurements associated with machine fatigue [38]. With the most commonly used acceleration type, the vibration condition is monitored by positioning it close to a bearing on the motor frame at high frequencies [38, 43].

Vibration analysis is used in many studies on mechanical failures also occurring in induction motors. Imbalance, misalignment, looseness, and bearing failures are specific signs in the vibration spectrum [41]. A condition monitoring strategy can be established by correlating certain fault types with specific frequencies, or by trend analysis with acceleration data. Depending on the application and user requirements, with vibration condition monitoring, a cost-effective maintenance plan can diagnose and take action before the machine and its components fail or cause performance loss.

2.3.3 Motor current monitoring

Condition monitoring via supply current signals provide useful information on not only for the motor itself but also the mechanical system that motor drive [44]. An important aspect of the maintenance strategy is the inclusion of a mechanical system that motor drives. Especially with VFD-fed systems, where current is already sensed to control motor operation, both electrical and mechanical faults can be diagnosed without additional sensor need [31, 44–46].

In industrial applications where ambient conditions are not suitable for vibration signal measurement, current monitoring may be preferred due to its robustness to ambient conditions, especially when disturbances are high. Therefore, current-based condition monitoring, which is proven in industrial applications, has benefits such as economical, versatile and reliable over other monitoring techniques.

Although many studies have been done to diagnose fault using current signals, studies on VFD-fed motors are limited. It should be noted that the PWM signals used in VFDs can suppress the characteristics of a failure in the motor current signal, making diagnosis difficult [47]. In this study, electrical and mechanical fault detection is emphasized in different load and frequency scenarios over the single-phase stator supply current of the VFD-fed three-phased induction motor.

2.4 Signal Processing Techniques

Signal processing, also named feature generation, can be defined as the extraction and interpretation of the characteristics of the sensor data received from the machine whose status is to be monitored [48]. In a sense, it is the transfer of expert knowledge to the system and its use in monitoring the motor condition. Signals such as vibration,

temperature and current are carried out to reach the information that is not always easily visible in the data, which is often required to be revealed [49]. Signals are generally studied in two different domains: time and frequency.

2.4.1 Time domain based signal analysis

Time-domain features may be beneficial to monitor the state of continuous dynamical systems [50]. The performance of the diagnosis is strictly dependent on the selection of features that represent the characteristics of the system. The selection of appropriate features, on the other hand, is based on expert knowledge to obtain a reliable and accurate diagnosis [51]. In practice, there exists a large range of indicators to reveal the system's state, but in this study statistical features such as Root Mean Square (RMS), Mean, Median, Standard Deviation, Kurtosis and Skewness are employed [52].

2.4.2 Statistical analysis

The main idea behind statistical analysis is to understand the location, which is the typical or central value of a data set, and variability, which is the spread of a data set according to centre and tails [53]. The mean and median values are used to find the location, while standard deviation indicates the spread. Skewness and kurtosis criteria can also be examined to better understand the data.

2.4.2.1 Mean

Commonly called average, is the sum of the samples in the dataset divided by the total number of samples [52]. The mean is one of the best indicators if the underlying distribution is normal, but lacks the robustness of validity [53]. That is, if the underlying distribution is not normal, mean-based confidence intervals tend to be imprecise.

$$\bar{Y} = \sum_{i=1}^N Y_i / N \quad (2.6)$$

where,

\bar{Y} is the mean;

N is the number of data samples;

2.4.2.2 Median

The median, which is the point in a dataset that is greater than half the numbers and less than the other half, tends to have a robustness of validity but not a robustness of efficiency [52, 53].

$$\tilde{Y} = Y_{(N+1)/2} \quad \text{if } N \text{ is odd} \quad (2.7)$$

$$\tilde{Y} = (Y_{N/2} + Y_{(N/2)+1}) / 2 \quad \text{if } N \text{ is even} \quad (2.8)$$

where,

\tilde{Y} is the median;

2.4.2.3 Root mean square

RMS is also known as quadratic mean and represents the magnitude of a varying signal [52, 54]. As one of the most applied feature for rotating machinery, especially in AC electric motors, it is used for roughly estimating motor load and detecting general noise level.

$$\text{RMS} = \sqrt{\frac{1}{N} \left[\sum_{i=1}^N (Y_i)^2 \right]} \quad (2.9)$$

2.4.2.4 Standard deviation

Standard deviation is the square-root of the variance which is arithmetic average of the squared distance from the mean [52]. Similar to mean, standard deviation is also one of the best estimator, but also suffers the same lack of precision in case of distribution is not normal [53].

$$STD = \sqrt{\sum_{i=1}^N (Y_i - \bar{Y})^2 / (N - 1)} \quad (2.10)$$

2.4.2.5 Kurtosis

Kurtosis a measure that is to be used to understand if the data peaked or flat relative to a normal distribution [52]. High kurtosis indicates that the dataset tends to have a prominent peak close to the mean, while the dataset with low kurtosis tends to have a flat peak close to the mean rather than a sharp peak [53].

$$Kurtosis = \frac{\sum_{i=1}^N (Y_i - \bar{Y})^4}{(N - 1)STD^4} \quad (2.11)$$

2.4.2.6 Skewness

Skewness represents a lack of symmetry in a data set. The dataset is symmetrical if it looks the same to the left and right of the centre point [52]. The left skew represents a negative value while showing that it is taller on the left than on the right [53]. The right skew indicates the opposite situation. The skewness of a symmetric dataset converges to zero, and it is zero for a normal distribution [53].

$$Skewness = \frac{\sum_{i=1}^N (Y_i - \bar{Y})^3}{(N - 1)STD^3} \quad (2.12)$$

2.4.3 Frequency based signal analysis

The frequency domain is needed to reveal properties of a signal that are hidden in the time domain. This need actually has different motivations. One of them, considering the operating conditions, industrial machines are quite susceptible to noise and disturbances [48, 55]. To suppress or eliminate these effects, frequency domain transformations are less expensive in terms of computational requirements than time-domain methods [48, 56]. Another motivation is that the fault characteristics can be seen better in the frequency spectrum, as it is widely applied in the literature [56].

Frequency domain analysis has different techniques including the Fourier transform of time-domain waveforms, but Fast Fourier Transform and Power Spectral Density methods will be examined in the thesis.

2.4.3.1 Shannon-Nyquist sampling theory

Hardware-wise, it is not possible to transfer an analogue signal to the digital environment as it is in the physical world. For this reason, sampling of the analogue signal is necessary in order to represent a signal digitally. The Shannon-Nyquist Sampling Theorem specifies conditions that must be satisfied in order for an analogue signal to be converted to a digital signal [57].

When an analog signal $x(t)$ is sampled with the period T_s , the resulting signal is the discrete signal $x_s(n \cdot T_s)$ with $n = 0, 1, 2, \dots$

Two condition must be satisfied for accurate representation of $x(t)$ [57]:

1. The frequency spectrum of $x(t)$ must be limited by some maximum frequency, such that f_{\max}
2. The sampling rate f_s , must be at least twice the maximum frequency f_{\max}

$$f_s \geq 2 \cdot f_{\max}$$

where, $f_s = \frac{1}{T_s}$

2.4.3.2 Fast Fourier transform

After an analogue signal is sampled and the shape of the signal is obtained, this signal is now in a form that can be processed and analyzed in the digital environment. As an example, in order to find the output expression of a linear and time-invariant system, the input function of this system in the time domain is multiplied by the pulse input response of this system and the resulting signal is integrated. This computationally cumbersome convolution operation becomes an algebraic multiplication in the frequency space [48]. Therefore, in signal processing studies, some transformation methods have been developed to find the equivalent of the signal in the frequency space. One of these methods is the Fourier Transform.

In practice, to calculate the frequency spectrum (frequency-amplitude expression of the Fourier Transform) of a signal, the Discrete Fourier Transform of the signal is calculated. The mathematical expression of DFT is as follows [55]:

$$X(k) = \sum_{n=0}^{N-1} x(n) \cdot \exp\left(\frac{-j2\pi nk}{N}\right) \quad (2.13)$$

where, $0 \leq k \leq N - 1$ and N is the number of samples in discrete signal.

Since the form of the Discrete Fourier Transform given by the equation 2.13 requires N^2 complex multiplication and $N \cdot (N - 1)$ addition in the computer environment, the computational load is quite high especially for large N [42, 55, 57, 58]. For this reason, some Fast Fourier Transform (FFT) algorithms have been developed for faster computation of DFT. Some practical aspects of FFT to be used in condition monitoring [49];

- Motor supply current sampling is usually done at 5 kHz. Therefore, the bandwidth of the sensor should be at least 10 kHz.

- Shannon-Nyquist theorem indicates that sampling frequency must be twice the maximum frequency, but in practice 10 times increases accuracy.
- Spectral resolution, $\Delta f = \frac{f_s}{N}$

2.4.3.3 Power spectral density estimation

The harmonics seen in the current spectrum resulted from the faults depend on the motor load and hence the slip. When the signal is processed with the FFT, it introduces errors as it averages the spectrum amplitudes over the sampling period [59, 60]. PSD, on the other hand, is more resistant to slip variations due to its ability to monitor different frequency bands. PSD estimation can be categorized into two technique: parametric and non-parametric [61, 62]. In the scope of the thesis, a non-parametric method, Welch's approach is investigated.

In Welch's method, the time domain signal is split into segments of a certain length with overlaps between its segments, and a time-domain window is applied to the individual data segments, then estimate PSD by computing DFT for each segment and finally, the calculated PSDs are averaged [5, 63, 64].

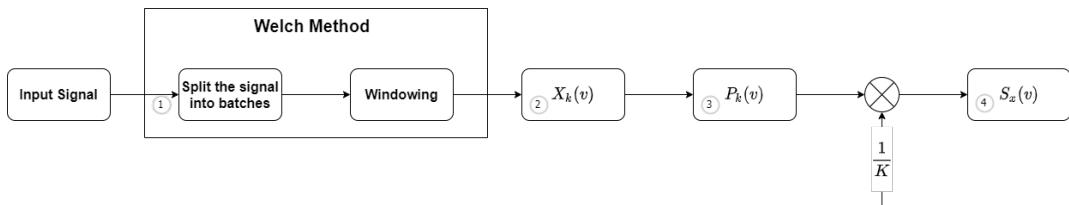


Figure 2.5 : Flowchart of Power Spectral Density estimation via Welch's method [5].

By segmenting and window overlapping the data, the Welch method can increase the resolution and also reduce both the variance and bias of the spectral estimation [61, 63, 65]. The mechanism of adding window overlaps on the signal, which is similar to noise removal with the recursion algorithms, results in a better Signal-to-Noise Ratio (SNR) in high noise data [5, 60, 66]. Welch's PSD estimation with Hamming window for fault diagnosis in induction motors outperforms FFT and periodogram methods in terms of robustness and accuracy due to reduced bias and variance [67].

For a given signal $x[0], x[1], \dots, x[N - 1]$;

1. Split the signal into K segments:

Segment 1: $x[0], x[1], \dots, x[M - 1]$

Segment 2: $x[S], x[S + 1], \dots, x[M + S - 1]$

Segment K : $x[N - M], x[N - M + 1], \dots, x[N - 1]$

where

M = Number of points in each segment

S = Number of points to shift between segments

K = Number of segments

2. For each segment ($k = 1$ to K), compute a windowed Discrete Fourier Transform (DFT) at some frequency $v = i/M$ with $-(M/2 - 1) \leq i \leq M/2$:

$$X_k(v) = \sum_m x[m]w[m] \exp(-j2\pi vm)$$

where

$$m = (k - 1)S, \dots, M + (k - 1)S - 1$$

w[m]= the window function

3. For each segment ($k = 1$ to K), compute the modified periodogram value, $P_k(f)$, from the Discrete Fourier Transform:

$$P_k(v) = \frac{1}{W} |X_k(v)|^2$$

where,

$$W = \sum_{m=0}^M w^2[m]$$

4. Average the periodogram values to obtain Welch's estimate of the PSD:

$$S_x(v) = \frac{1}{K} \sum_{k=1}^K P_k(v)$$

Figure 2.6 : Algorithm of Power Spectral Density estimation via Welch's method [6].

2.5 Data-driven fault diagnosis techniques

Model and signal-based techniques have been applied successfully for many years in condition monitoring and fault diagnosis studies in induction motors. Although these approaches have their own advantages, they require a certain level of field expertise. With the establishment of the Industry 4.0 phenomenon, the increasing data size in industrial applications in recent years and the developments in the technologies of the hardware that will store and process this data form the infrastructure for data-oriented approaches. The fact that there has been vast academic interest in machine learning techniques, which are also defined as artificial intelligence, especially in the last ten years, causes an increase in industrial applications.

Increasing data-driven studies in induction motors will be examined under machine learning and deep learning methods in this study.

2.5.1 Machine learning methods

The state information of the motor contained in the current signal usually requires the processing of the signal. Fault diagnosis can be made with machine learning methods by processing the signal with statistical approaches in the time and frequency domains and approaches to extracting characteristics at certain frequencies in the frequency domain.

Machine learning methods can be divided into supervised, unsupervised and semi-supervised according to training approaches. In supervised learning, the model is trained by giving the class information about the data to be processed as input, and the predicted class is given as a certain probability value as the model output. Within the scope of the thesis, supervised machine learning models were trained with a label indicating the motor status with the current signal.

2.5.1.1 Naive Bayes

Naive Bayes is a classification method based on the assumption that each feature is independent according to the class parameter [68]. Outputs a probabilistic prediction of the target class for an unknown observation with its characteristics and class

information input [69]. Although it is simple in structure, it can be considered as a benchmark model in classification studies because it gives very fast and effective results even when the independence assumption is violated [70].

2.5.1.2 k-Nearest neighbors

One of the simplest machine learning methods, kNN can give very good results despite its simple structure and easy application. Even if the data set is large, the nearest neighbors can be found quickly and classification can be made [71].

It works on the assumption that observations with similar attributes are more likely to produce similar results [72]. As a non-parametric method kNN performs classification based on the number or distance of the Nearest Neighbor in the data set.

2.5.1.3 Support vector machines

SVM as a supervised machine learning method is widely applied in regression problems and generally in classification problems. The main concept of SVM is locating a hyperplane that separates and classifies the datasets. The margin defines the support vectors of the datasets as it is the closest point to the hyperplane. Classification is accomplished by finding the hyperplane between the classes.

Besides performing linear classification, SVMs can efficiently perform nonlinear classification using various kernel functions. Kernel selection is quite important as it directly affects classification performance [73]. There are many kernel and SVM architecture options for various problems in the literature. On the other hand, the performance of SVM algorithms is limited as the computational load increases with the increase in the number of samples [74].

2.5.1.4 Multi layer perceptron

MLP is basically a learning structure consisting of input, output and hidden layers. The input layer is the layer where the input signals arrive and the output layer is the layer where the output signals depart. The layers other than input and output layers are called hidden layers. All layers in MLP are made up of neurons. These parts are called neurons because the structure of the brain's neurons to receive input and transmit an output by processing this input is imitated. In MLP, the output of each layer becomes

the input of the next layer. The inputs to each neuron are multiplied by a weight and summed, and a weight (bias) is added to the value found. The value reached is passed through an activation function while being transmitted to the next layer. These processes are repeated from the input layer to the output layer. This process is the forward propagation part of MLP [74]. Initially, the feed-forward stage takes place with randomly determined weights and an estimated output value is calculated for each signal. Each estimated output value calculated in this way contains errors. The weights are updated according to the errors to try to reduce the error. This process is performed starting from the output layer to the input layer. This part of the MLP is called backpropagation [74]. During this update process, how much the weights will be updated is determined by the learning rate. The number of times these operations will be repeated for all data is expressed as the epoch. For how much data forward propagation and backpropagation will be applied together is represented by the batch size.

2.5.1.5 Ensemble learning

Each model is developed based on certain assumptions and errors that occur as a result of these assumptions. It is aimed to minimize the total error by combining the predictions with a kind of voting method between training multiple models and their outputs [75, 76]. The practice of employing multiple classification models and combining their predictions in this way is called ensemble learning.

Random forest

Bagging, which is a voting method, can be defined as creating different variations by changing a dataset slightly and combining the outputs of each model after being trained on a different variation [75, 76]. As a refinement of bagged trees, Random Forest mainly aims to improve bagging by decorrelating the trees [77]. The procedure of voting for the most popular class after a large number of trees has been created is called random forests [77].

XGBoost

The most common ensemble method, boosting provides sequential learning of the predictors and can be considered as a model averaging [76]. As an iterative process, it continues to add classifier learner until a limit is reached by means of the number

of models or accuracy [75]. XGBoost, a scalable machine learning algorithm for large-scale tree boosting, is a classifier variant with high performance in many different applications [78].

2.5.2 Deep learning methods

Deep learning, which is a sub-branch of machine learning, as its name implies, is artificial neural networks that deepen using many hidden layers. In machine learning, the artificial intelligence model requires feature engineering as a kind of preprocessing. Feature engineering requires specialized knowledge of statistical computing and signal processing, as well as knowledge of the general characteristics of the system. Deep learning methods, which provide a direct link between data collection and decision output by eliminating the manual feature extraction process, have attracted great attention in recent years.

As the importance of data spreading with Industry 4.0, manufacturers are starting to offer add-ons and services that will make it easier to access data for their products. For example, Wat Motor now offers special designs for vibration and temperature sensors that will be mounted into the motor, allowing monitoring of the status of its induction motors [79]. In the coming years, with the increase in the available data, deep learning methods will attract even more attention. While the performance of deep learning techniques increases in parallel to available data amount compared to machine learning methods, the trained model can be transferred to other applications with its transferability feature [80].

The increase in accessible data day by day, the interest of researchers from industry and academia that lead the development of new methods and approaches, and hardware developments such as Graphics Processing Unit (GPU), Tensor Processing Unit (TPU) and Field Programmable Gate Array (FPGA) that can process excessive calculations are the driving force of deep learning [80, 81].

2.5.2.1 1D-Convolutional neural networks

CNN is a special case of feed-forward neural networks, with each hidden layer being a convolutional and pooling layer [82]. Thanks to the convolutional layer, it provides a serious advantage in computational load [82]. CNNs differ from Neural Networks

in that they extract features of higher-order features from the input with convolutional processes [78]. The pooling layer is a layer that works in the direction of this idea by downsampling.

It is generally used in image processing and relies on the idea that an image does not need to be seen in its entirety to be perceived [83]. Although it was originally a 2D structure, CNN structures of different dimensions were built afterwards and especially 1D structures employed in the analysis of time series signals such as current and vibration signals [84].

2.5.2.2 Long-Short term memory networks

Recurrent Neural Network (RNN) is a special type of neural network for sequential data. RNN also uses information from previous time steps, so it remembers the entire sequence [85]. In the backpropagation process for RNN, some problems may occur because the gradients are multiplied. These problems can be observed as the gradient becomes too small or too large [85]. To avoid these situations, LSTM, which has a memory cell and a special case of RNN, is used. Memory cell consists of forget, input and output gates. The forget gate determines how much previous step information should be remembered. The input gate expresses how much of the newly acquired information needs to be remembered. The output gate indicates how much memory will be used for the output of the relevant time step [86].

2.6 Performance evaluation

Various metrics, which can be aggregated under binary or multi-class classification, are used to compare the performance of algorithms to be used in fault diagnosis [87,88]. In the diagnosis of asynchronous machine or their components, binary classification can be made as healthy or faulty condition, while multiple classification metrics should be used when separating two or more fault types. Although there is no definite consensus for the metrics used in the comparison between the classification methods, the most frequently used metrics for multi-class classification will be examined in this section.

In order to create metrics, certain measures must be introduced. The confusion matrix shows the actual and predicted classification using certain measures. In the motor diagnostics specific, these four metrics can be defined as follows:

		Predicted Class	
		Positive	Negative
Actual Class	Positive	True Positive	False Negative
	Negative	False Positive	True Negative

Figure 2.7 : An example of Confusion Matrix.

True Positive (TP), the state where both the actual and predicted values are healthy,
 False Positive (FP), the classification of the actually faulty condition as faulty,
 False Negative (FN), the classification of the actually healthy motor as faulty,
 True Negative (TN), the state where both the actual and predicted values are faulty.

2.6.1 Precision and recall

Precision refers to the ratio of samples that the classification method predicts as healthy to actually healthy samples, and shows how reliable the healthy motor prediction can be. Recall, on the other hand, expresses how many of the healthy motor samples were labelled correctly as a result of the classification and shows the model's ability to find the healthy motor.

$$Precision = \frac{TP}{TP+FP} \quad (2.14)$$

$$Recall = \frac{TP}{TP+FN} \quad (2.15)$$

2.6.2 Accuracy

Accuracy, one of the most common model performance metrics, is an indicator of how well it can distinguish between healthy and faulty motors in the entire data set [89]. In general, the number of healthy state data is naturally higher in diagnostic applications that result in an unbalanced data set [90]. In such a case, evaluating only with the accuracy metric may lead to catastrophic situations.

$$Accuracy = \frac{TP+TN}{TP+FP+FN+TN} \quad (2.16)$$

2.6.3 F-measure

As a metric based on Recall and Precision, F-Measure allows better inference than accuracy on classification performance, especially on unbalanced datasets [91]. In multi-class classification case F-measure needs to be modified as Macro F-measure by averaging each and every class' F-measure.

$$F\text{-measure} = 2 \cdot \frac{\text{Recall} \cdot \text{Precision}}{\text{Recall} + \text{Precision}} \quad (2.17)$$

$$\text{Macro F-measure} = \frac{1}{C} \cdot \sum_{c_i \in C} \text{F-measure}(c_i) \quad (2.18)$$

where:

c_i is the reference class

C is the total number of classes

Especially in condition monitoring applications, fault alarms even though there is no fault increase the maintenance cost and at the same time, it can disrupt the operation. On the other hand, missing a fault condition can also damage equipment and disrupt the operation. The performance of the classification method is important in terms of optimizing both cases. F-measure can respond to this optimization as it contains components for these two states [88, 92].

2.6.4 Cohen's kappa

Another metric that works well with unbalanced data, Cohen's Kappa correlates the estimated and actual values, taking into account the imbalance in the class distribution [93]. By removing the random dependency between the predicted and actual classification, enables to compare different classifiers [89].

$$\text{Cohen's Kappa} = \frac{d \cdot n - \sum_k^C p_k \cdot t_k}{n^2 - \sum_k^C p_k \cdot t_k} \quad (2.19)$$

where:

C is the total number classes

$d = \sum_k^K C_{kk}$ the total number of correctly predicted samples

$n = \sum_i^K \sum_j^K C_{ij}$ the total number of samples

$$p_k = \sum_i^K C_{ki} \quad \text{the number of instances that class } k \text{ was predicted (column total)}$$

$$t_k = \sum_i^K C_{ik} \quad \text{the number of instances that class } k \text{ actually presents (row total)}$$

2.6.5 Area under the curve

The receiver operating characteristic (ROC) curve, which is another method widely used in binary classification performance measurement, shows the performance of classification methods in two dimensions [94]. To reduce this metric to one dimension, the area under the ROC curve (AUC) is calculated.

In multiclass classification problems, AUC values now transform into multiple binary classification values. The specified formula is used to reduce to a single numerical value [94]:

$$AUC_{total} = \sum_{c_i \in C} AUC(c_i) \cdot p(c_i) \quad (2.20)$$

where:

c_i is the reference class

C is the total number of classes

$p(c_i)$ is the prevalence of the reference class in the dataset

$AUC(c_i)$ is the area under the class reference *ROC* curve for c_i

According to this formula, AUC values are calculated by creating a ROC curve for each reference class, and the AUC_{total} value is obtained by weighting it with the prevalence of the reference class [91, 94]. The advantage of this method is that it is easily computable and is derived directly from reference class ROCs [94].

3. EXPERIMENTAL SETUP AND METHODOLOGY

This chapter is devoted to explaining the experimental setup and the data obtained for the condition monitoring and diagnostics of the induction motor. Experiments were carried out in the General Purpose Industrial Motor Laboratory at WAT Motor Company facilities. During the studies, the faults related to the bearing, stator and rotor of the squirrel-cage induction motor were artificially created and their data were collected. The test system used in the study is shown in Figure 3.1.



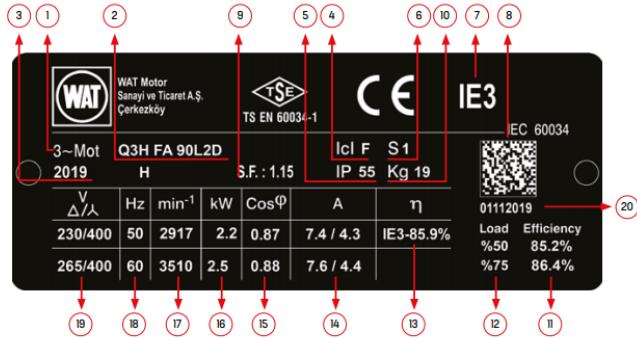
Figure 3.1 : Experiment setup, courtesy of WAT Motor Co.

Table 3.1 : Equipments used in experimental-setup.

Equipment
Kistler 200 kW Dynamometer
Yokogawa Power Analyzer
Teledyne LeCroy Oscilloscope
WAT-11 kW Induction Motors
WAT-WF-80 General Purpose Variable Frequency Driver

In the experimental studies, the stability of the system was tracked by monitoring the three-phase current and voltage signals on the power analyzer. One phase of the stator supply current of the motor was recorded with a current probe at a sampling frequency of 5 kHz for 10 seconds. In addition, the load torque acting on the motor and the rotor rotation speed of the motor were measured via the dynamometer system.

Experimental studies were carried out at a rated load of 72 Nm and a frequency of 30, 35, 40, 45 and 50 Hz at 54 Nm, which is 75% of the rated load, by applying the



- | | | | |
|----|---|----|--|
| 1 | Motor type | 11 | Efficiency value (acc. to IEC 60034-2-1) |
| 2 | Motor code | 12 | Load value |
| 3 | Year of manufacture | 13 | Efficiency value (acc. to IEC 60034-2-1) |
| 4 | Insulation class | 14 | Nominal current |
| 5 | IP class | 15 | Power factor |
| 6 | Service type | 16 | Motor output power |
| 7 | Efficiency class (acc. to IEC 60034-30) | 17 | Rated speed |
| 8 | 2D Barcode | 18 | Motor nominal frequency |
| 9 | Service factor * | 19 | Operation voltage |
| 10 | Motor weight | 20 | Production tracing number |

Figure 3.2 : Typical induction motor label, courtesy of WAT Motor Co.

Table 3.2 : Nominal values of WAT Motor 3-phase induction motor.

Voltage (V)	Power (kW)	Power (Hz)	Speed (rpm)	Current (A)	Torque (N · m)
400/690	11	50	1475	22.0/12.7	71.3

v/f control method over the VFD. Three motors produced in the same series on the same production line were taken and their data were collected in "healthy" condition under the mentioned conditions. For bearing failure, a fault was artificially created after hammering the drive-end bearing of one of the motors, while for stator failure, the insulation between the two turns in one phase of another motor was eroded. For the last motor, one of the bars in the rotor cage has been drilled for rotor failure. It should also be noted that studies conducted only for two cases in steady-state conditions for the motor as healthy and faulty.

Table 3.3 : Brief information about experimental conditions.

Fault Type	Frequency (Hz)	Load (N · m)
Bearing	30	≈ 72
Stator turn-turn	35	≈ 54
Broken Rotor Bar	40	
	45	
	50	

Another issue to be considered while collecting data is the amount of data required to see the effects of failure. In the experiments, the mechanical speed of the rotor varies

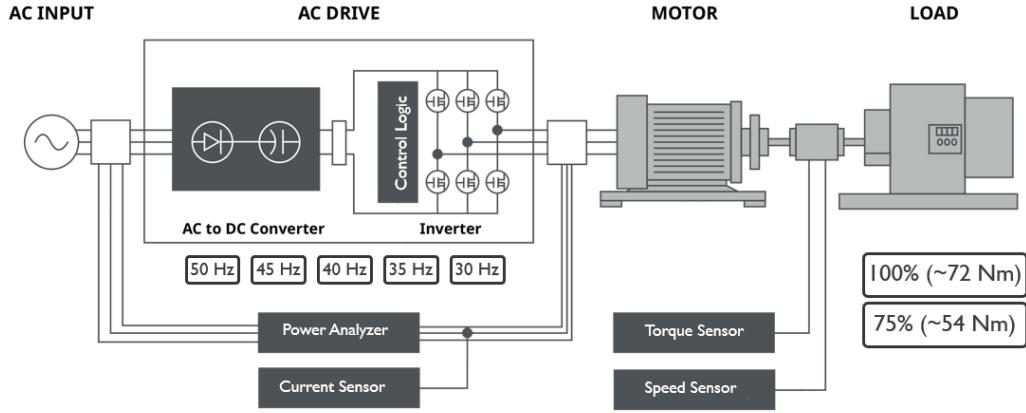


Figure 3.3 : Schematic of test system.

between 857 and 1472 rpm in proportion to the reference control command applied. According to the equation given in 3.1 [86], approximately 200 data points are taken per revolution of the rotor shaft with 5 kHz sampling frequency and sampling time of 10 seconds, a window width of 50,000 data points, fault impacts expected to be easily captured.

$$\text{Number of Data Points} = \frac{60 \cdot \text{Sampling Frequency (Hz)}}{\text{Rotor's Mechanical Speed}} \quad (3.1)$$

In data analysis, studies were carried out in time and frequency domains. In the time domain, the stator supply currents are studied in raw (no pre-processing), while in the frequency domain, Welch's power spectral density estimation is applied to the current data. Current signals obtained in 1 second for healthy and faulty conditions under different speed and load scenarios, as well as 10 times zoomed, are shared in Figure 3.4 to Figure 3.9. Looking at the plots, it is not possible to interpret whether the motor is faulty in the time domain or even if it is faulty, which fault has occurred.

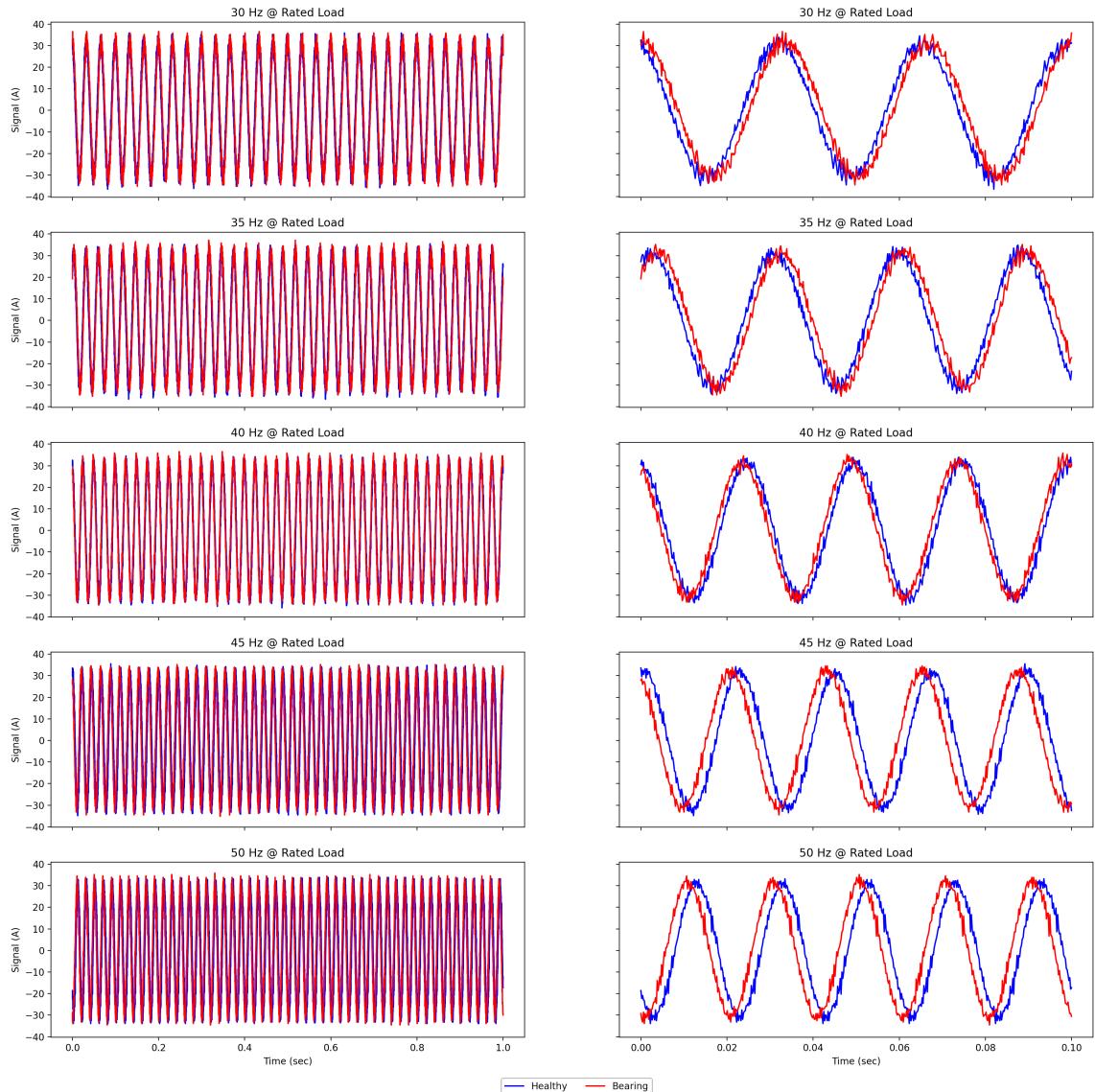


Figure 3.4 : An example of stator current signals of the healthy and bearing-fault motor at rated load.

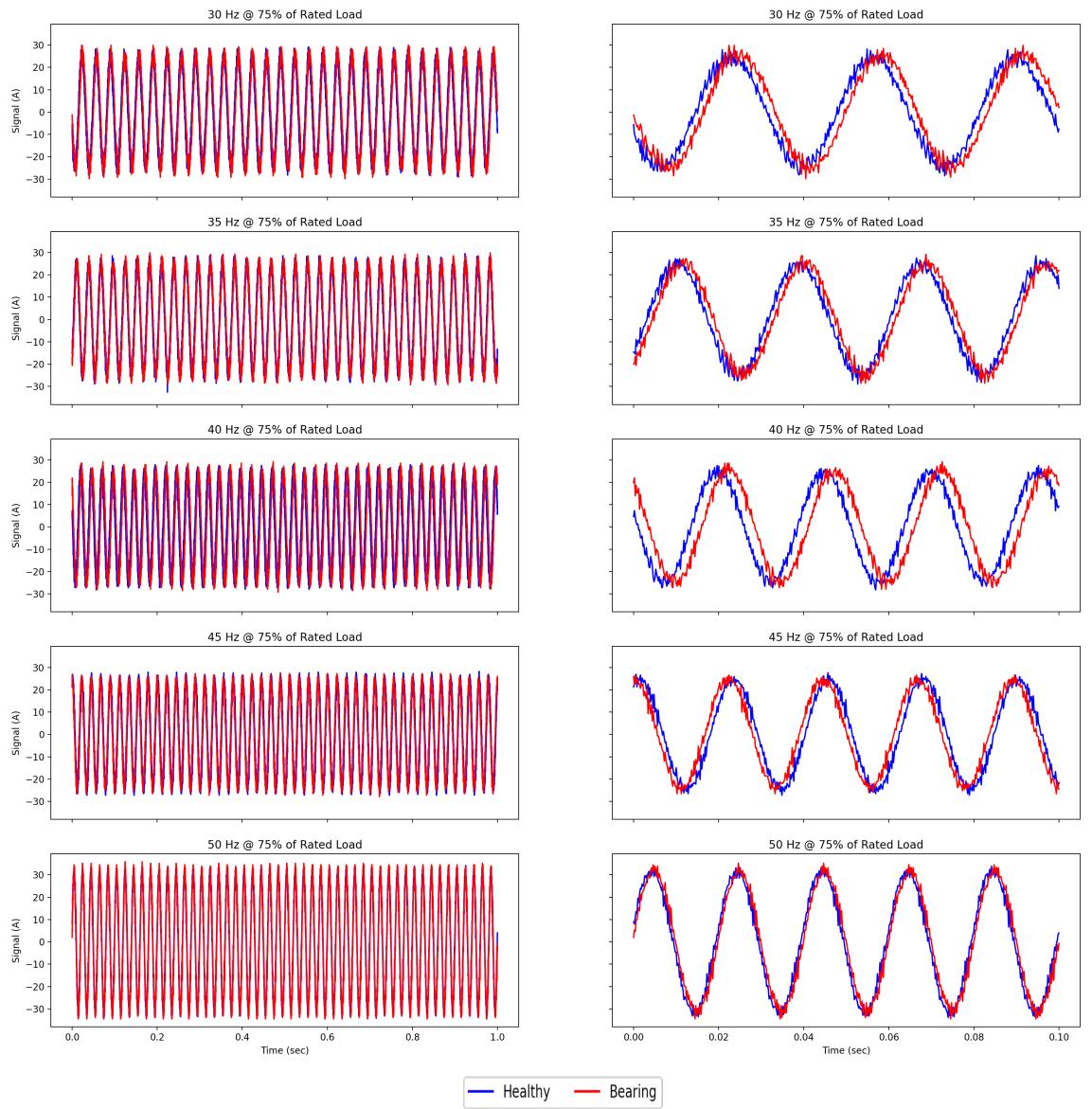


Figure 3.5 : An example of stator current signals of the healthy and bearing-fault motor at 75% of the rated load.

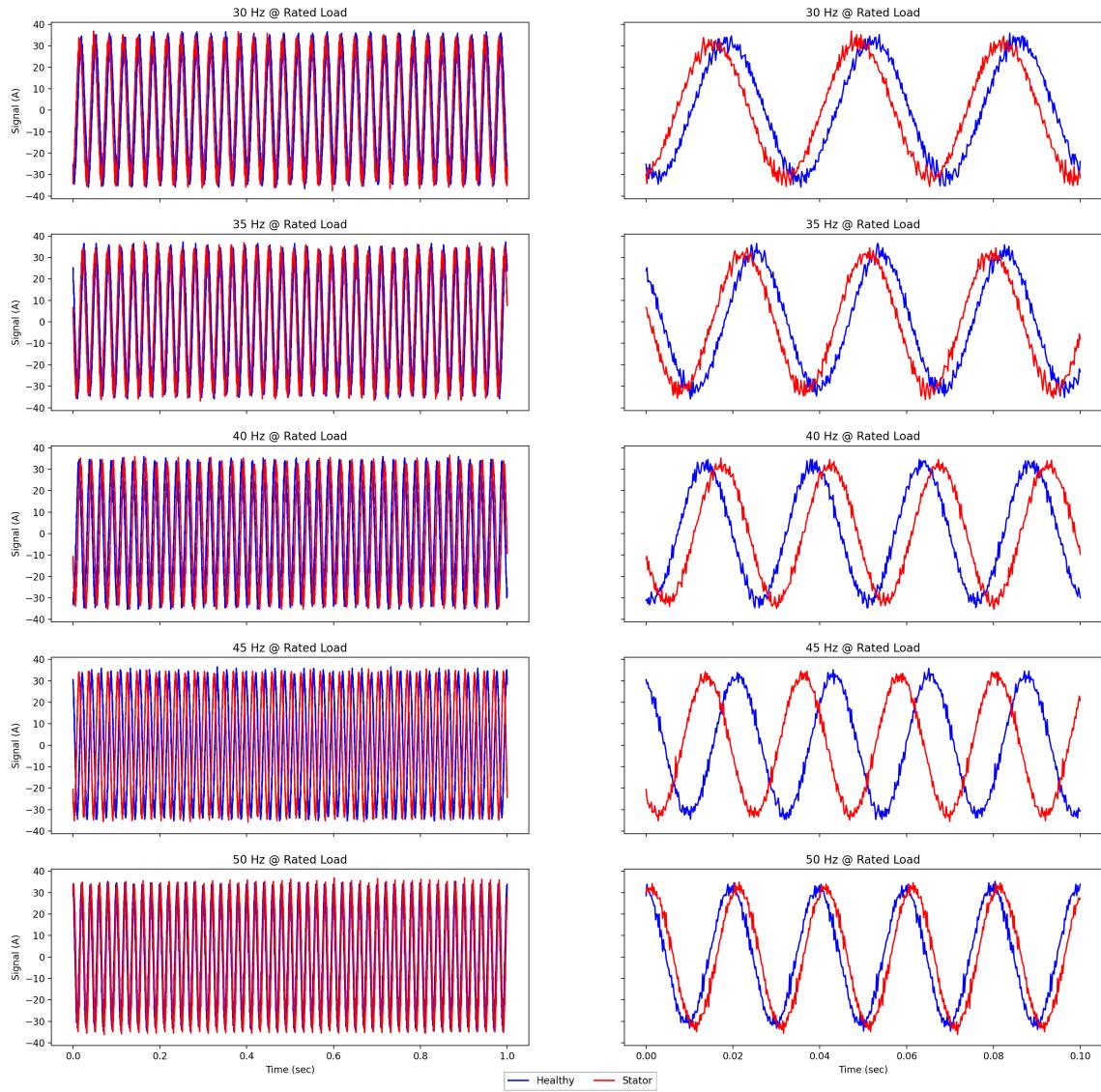


Figure 3.6 : An example of stator current signals of healthy and stator inter-turn-fault motor at 75% of the rated load.

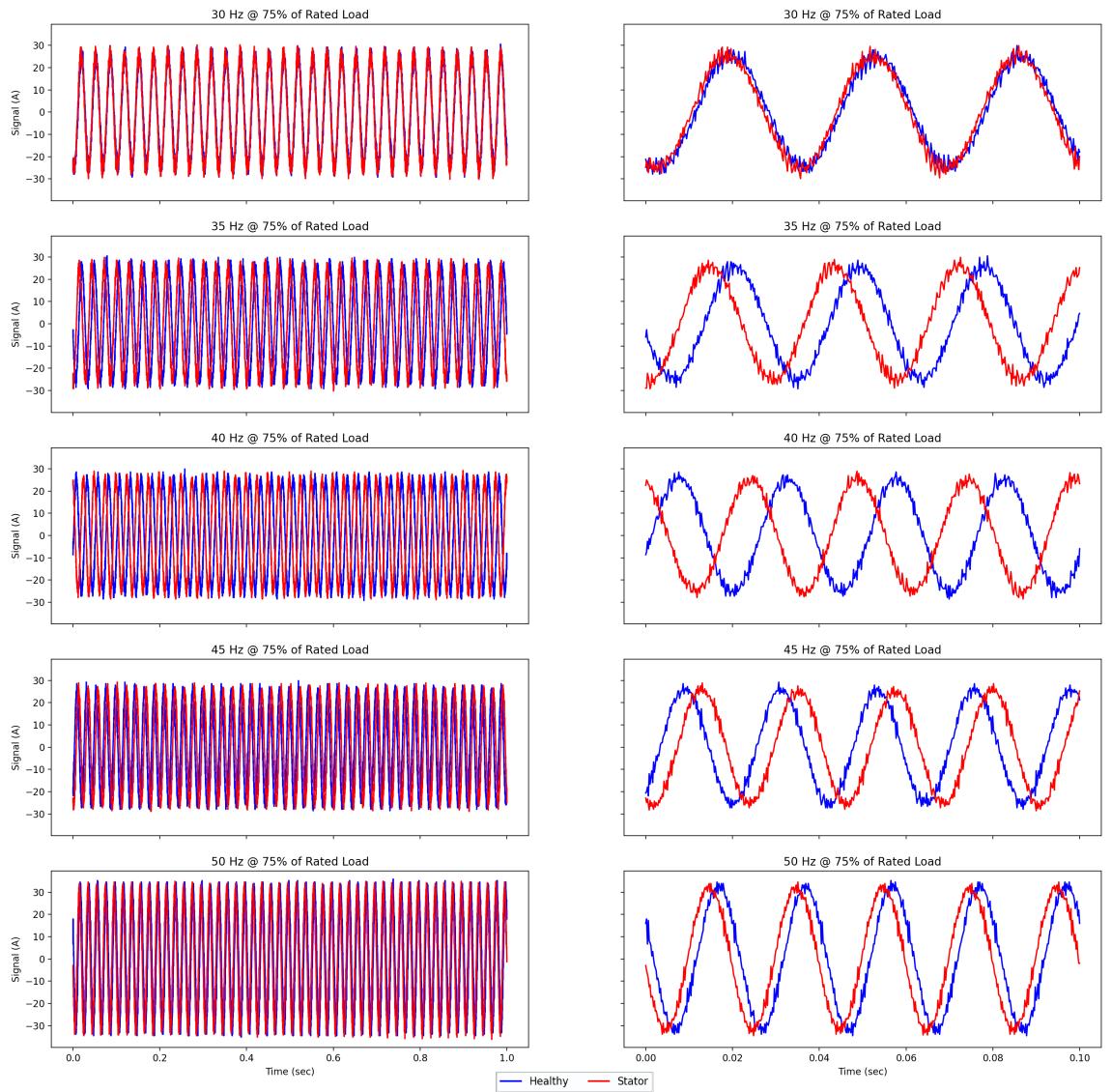


Figure 3.7 : An example of stator current signals of healthy and stator inter-turn-fault motor at 75% of the rated load.

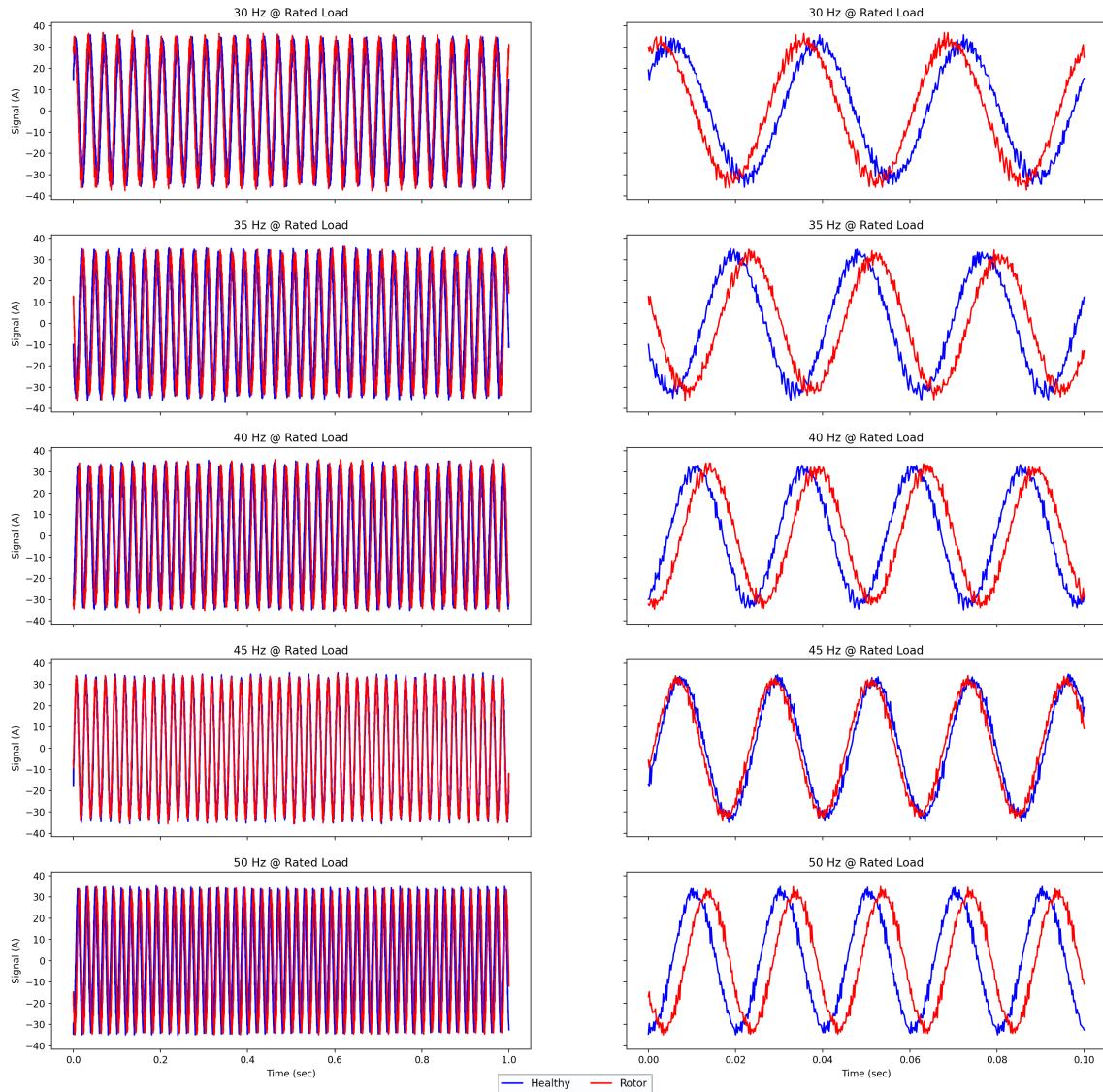


Figure 3.8 : An example of stator current signals of healthy and broken rotor bar-fault motor at 75% of the rated load.

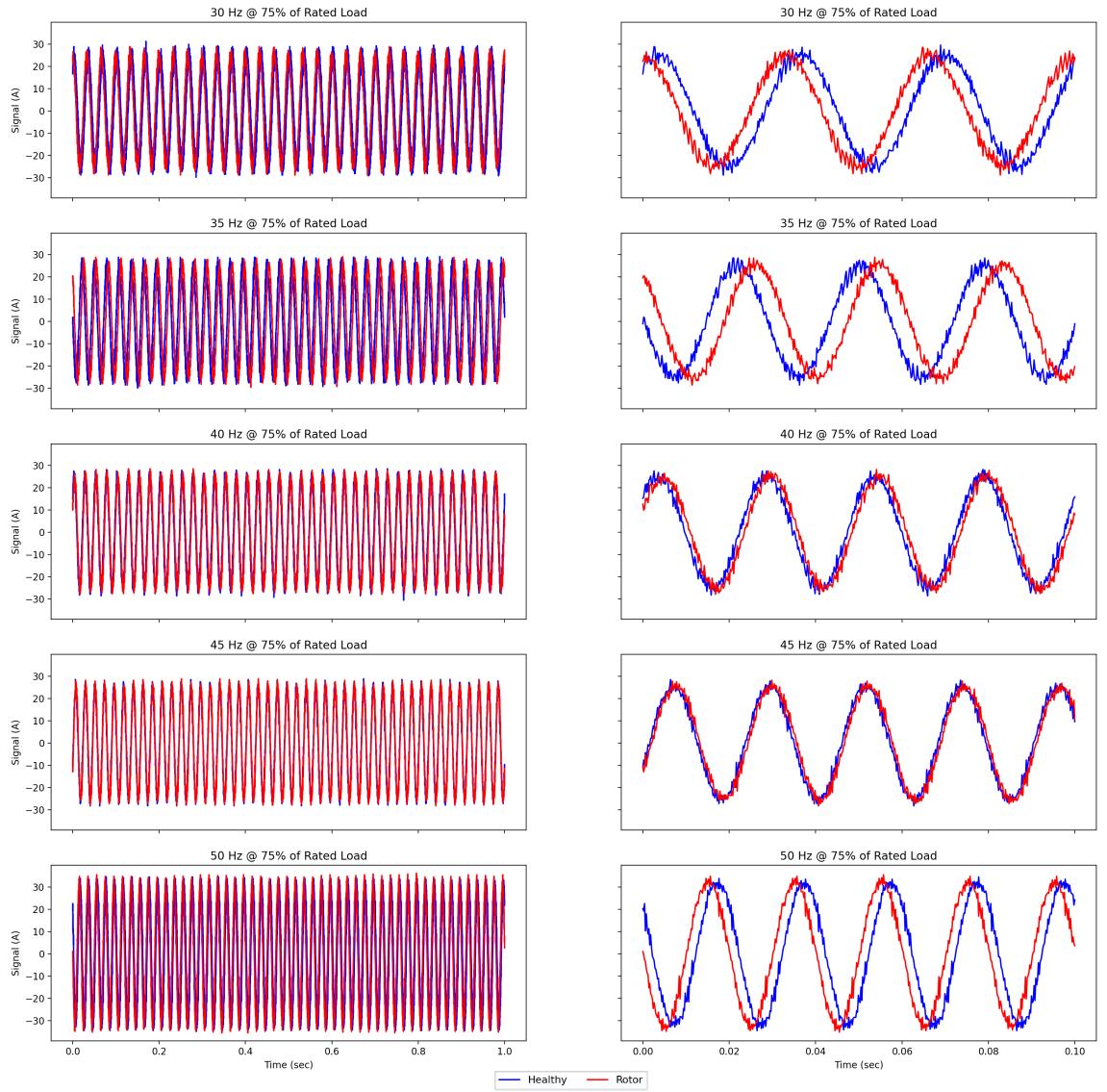


Figure 3.9 : An example of stator current signals of healthy and broken rotor bar-fault motor at 75% of the rated load.

In addition, Figures 3.10 to 3.15 demonstrate Welch's PSD estimation applied to healthy and faulty conditions of stator currents. Looking at the PSD graphs, there are visible differences between the faulty states and the healthy states, as well as the variation created by different faults at certain frequencies.

Table 3.4 : Input parameters for estimating Welch's PSD.

Parameter	Value
Window Type	Hamming
Overlap (%)	50
Number of DFT Points	50000
Sample Rate	5000

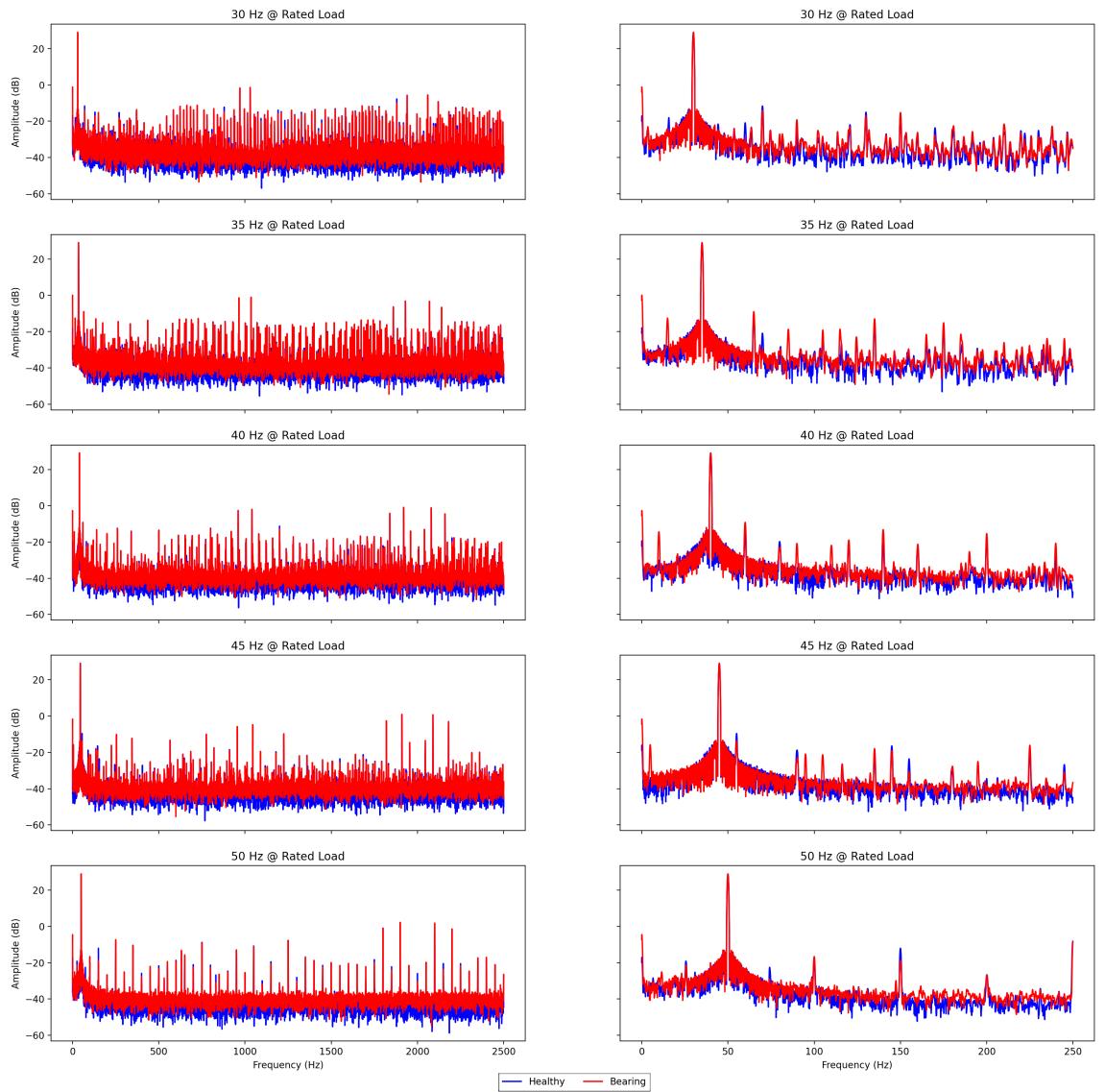


Figure 3.10 : Welch's PSD estimations of the healthy and bearing-fault motor at rated load.

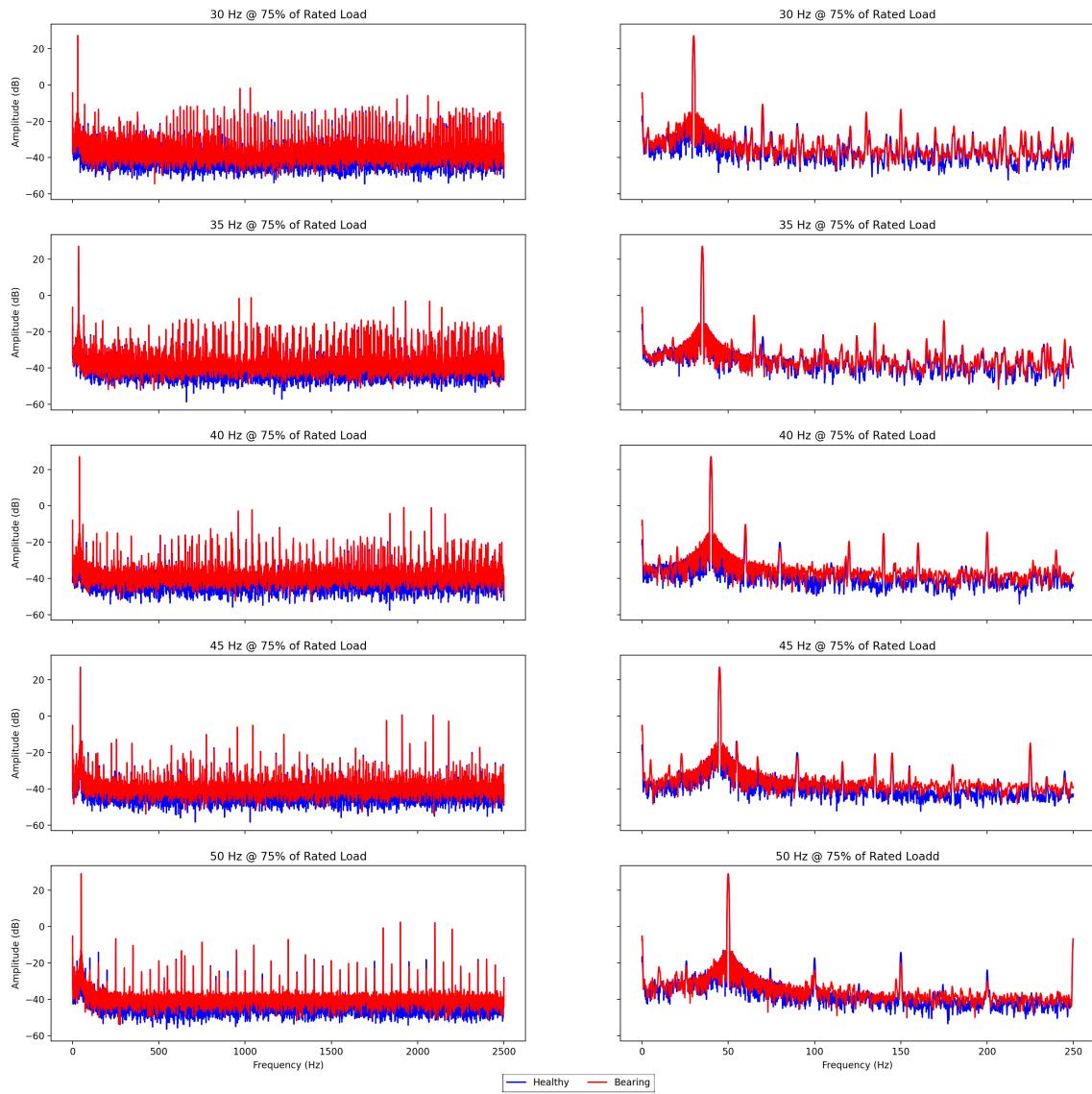


Figure 3.11 : Welch's PSD estimations of the healthy and bearing-fault motor at 75% of the rated load.

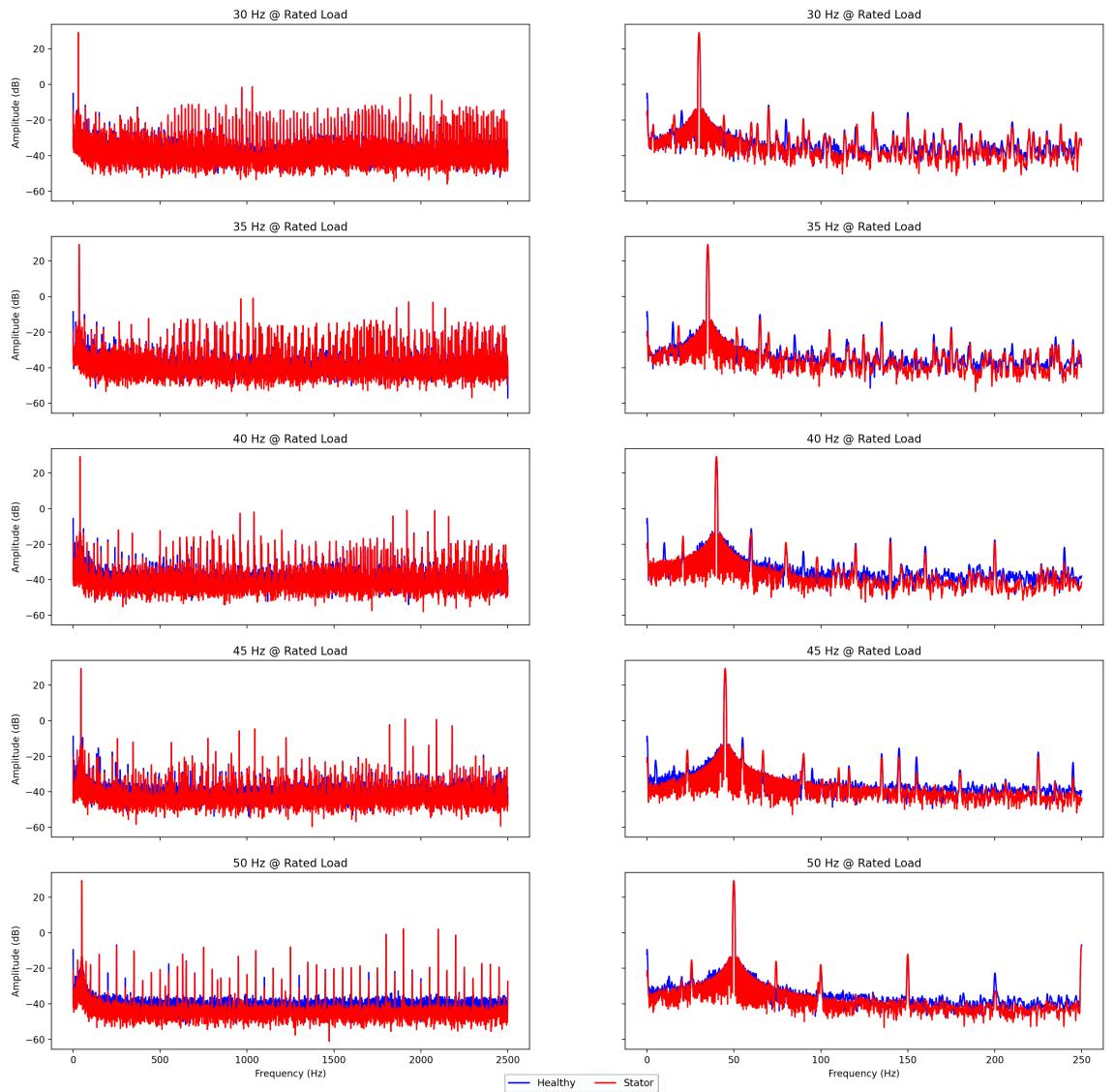


Figure 3.12 : Welch's PSD estimations of the healthy and stator inter-turn-fault motor at 75% of the rated load.

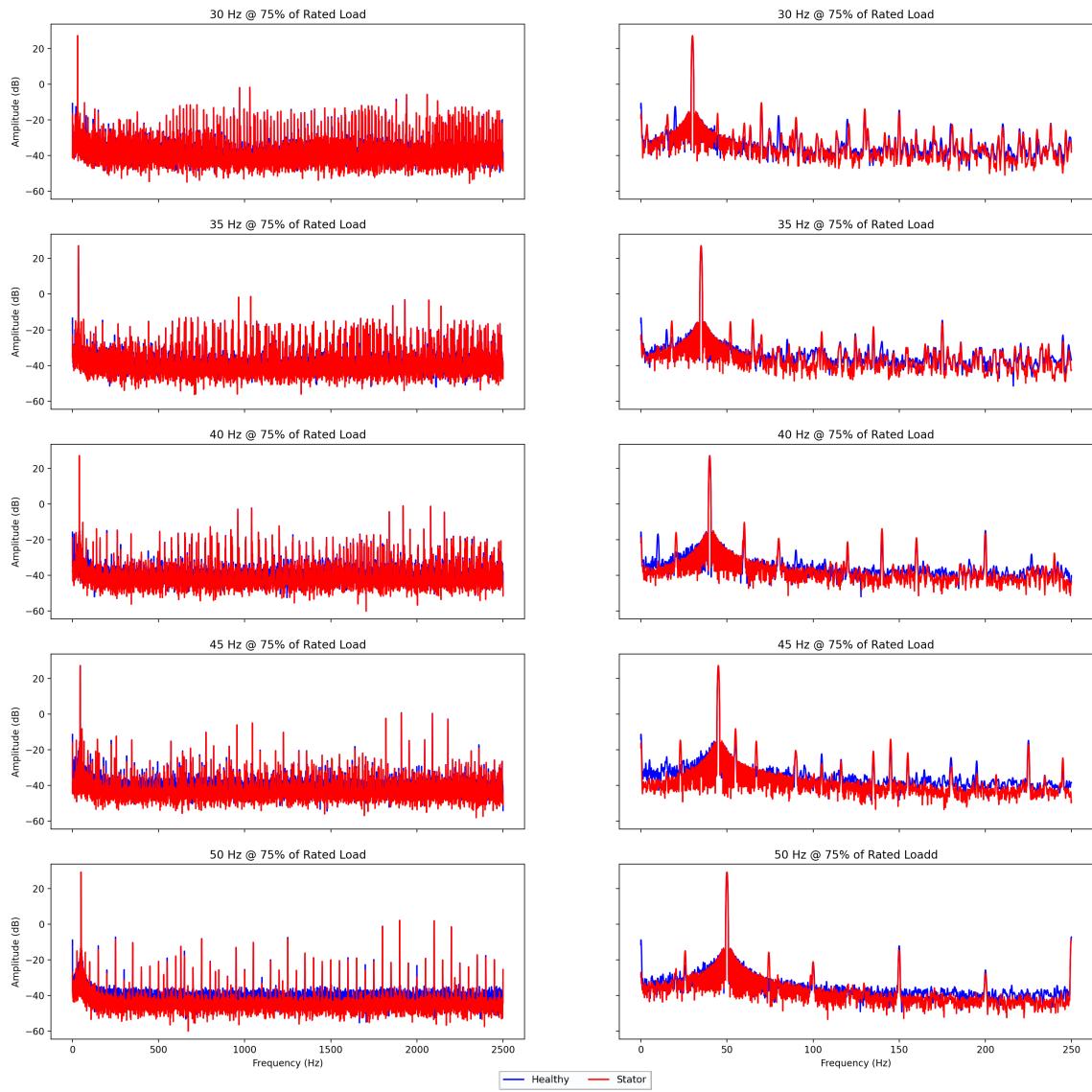


Figure 3.13 : Welch's PSD estimations of the healthy and stator inter-turn-fault motor at 75% of the rated load.

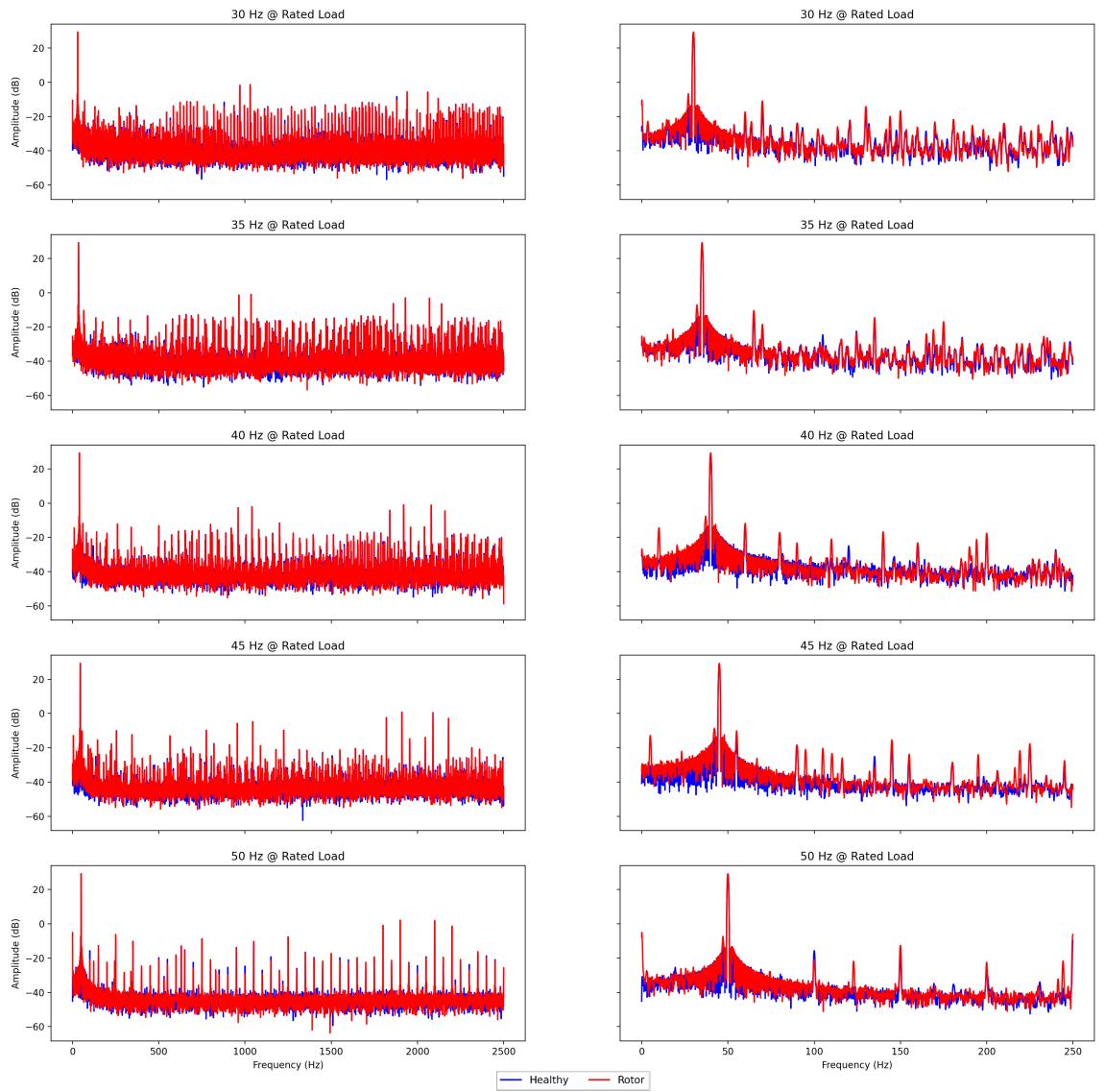


Figure 3.14 : Welch's PSD estimations of the healthy and broken rotor bar-fault motor at 75% of the rated load.

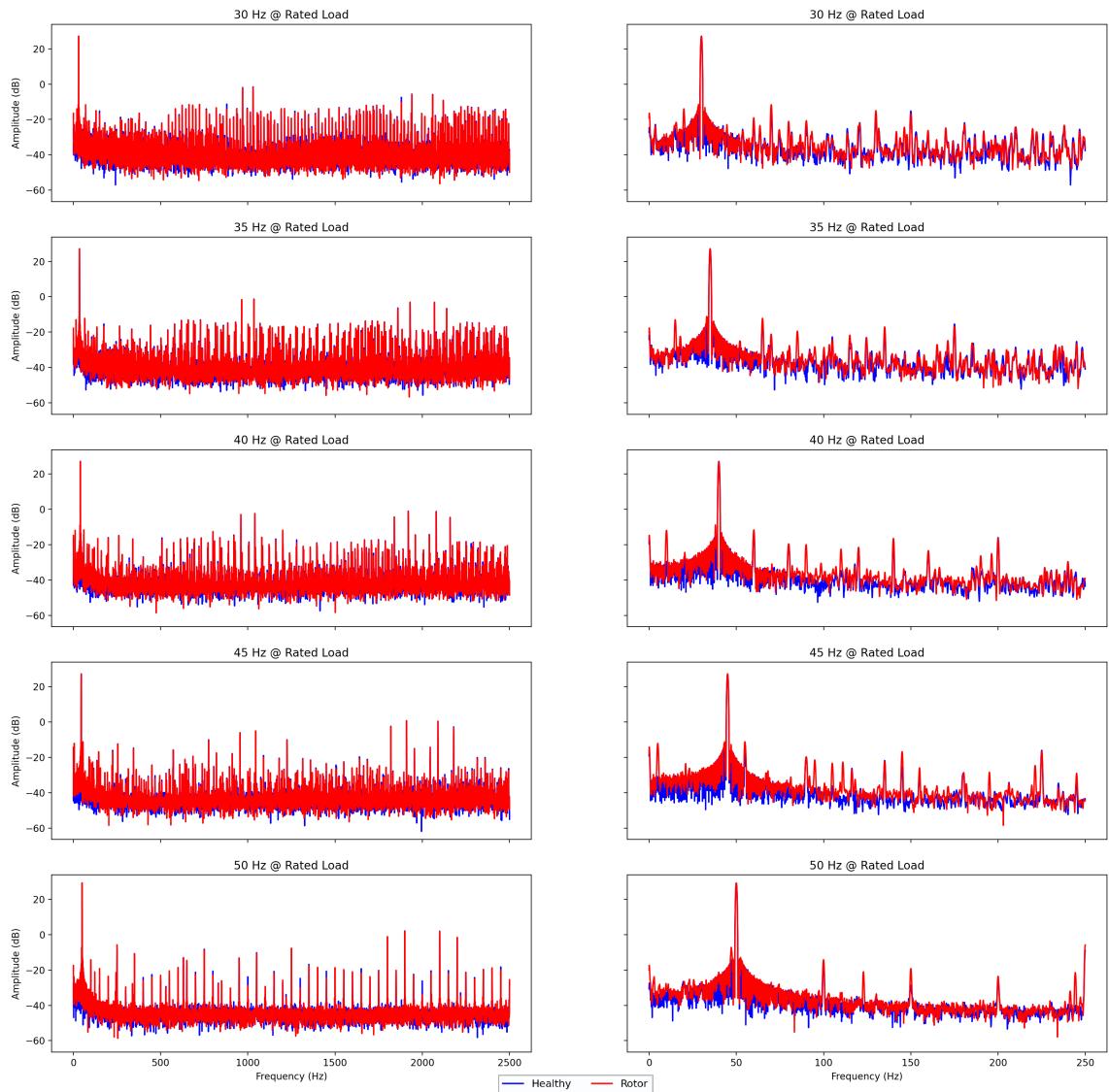


Figure 3.15 : Welch's PSD estimations of the healthy and broken rotor bar-fault motor at 75% of the rated load.

4. FAULT DIAGNOSIS METHODOLOGY

This chapter discusses four different methods for performing fault diagnosis. First of all, different signal processing methods are applied to the 1-phase current signal and compared with machine learning and deep learning methods. In the context of the thesis, two different comparisons are made. The first is to examine the advantages and disadvantages of different feature extraction scenarios and compare them with various metrics, while the second is to examine deep learning methods as an end-to-end solution.

4.1 Machine learning analysis

Machine learning methods need preprocessing for feature extraction, albeit in different ways. These features, which will be utilised for fault detection, directly affect the performance of the classifier. Within the scope of this thesis, three different signal processing for machine learning methods and the calculated features based on this process will be examined. In addition, the responses of the VFD powered motor to different speed and load scenarios are also covered.

4.1.1 Time domain statistical analysis

The first method is to extract the statistical properties of the current data in the time domain and to diagnose faults with machine learning techniques over them. Statistical features such as kurtosis, skewness, mean, RMS, standard deviation and median are extracted in the Knime to be used as features to predict whether the motor is healthy or faulty.

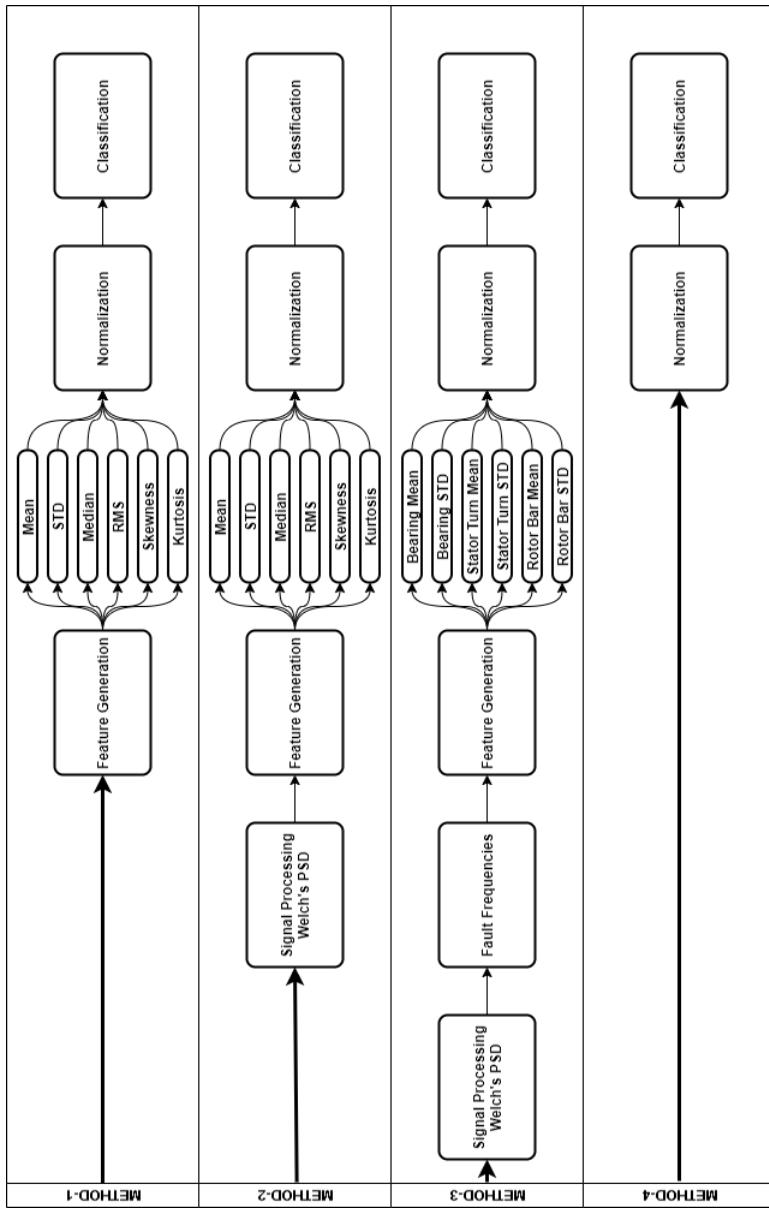


Figure 4.1 : Flowcharts of the methods presented in this thesis.

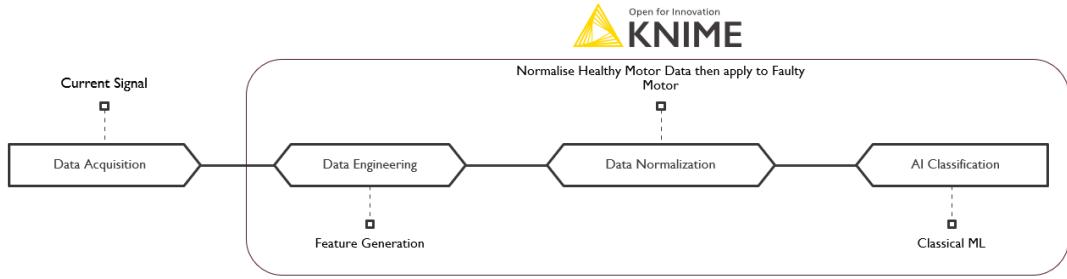


Figure 4.2 : Diagram of time-domain statistical analysis method.

In order to examine the distributions of these features in different classes, t-SNE, a manifold learning technique that reduces high-dimensional data to two or three dimensions, was used [95]. It is widely preferred because it can capture nonlinear structures by exploiting local relationships between data points. It is employed to reduce the 6-dimensional feature space to 3 dimensions. It seems that a short-circuit fault between the stator windings and bearing fault can be differentiated from the healthy condition, while a 1-bar broken rotor fault is relatively more difficult to differentiate. This situation can be understood as it does not change the motor behaviour for 1-bar broken rotor failure, but when classification is made with machine learning methods, it is seen that there are algorithms with high performance.

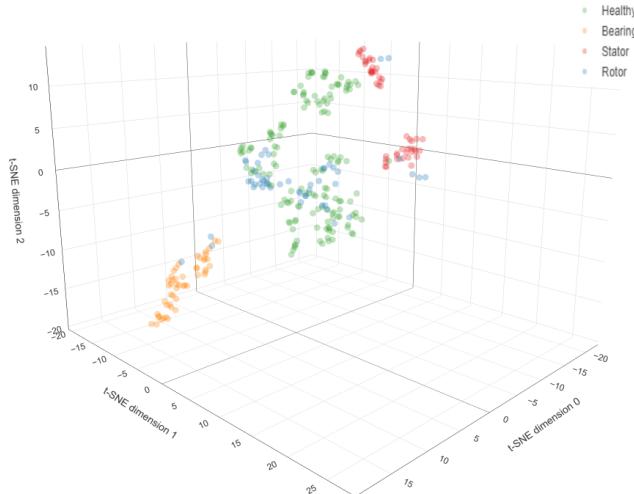


Figure 4.3 : t-SNE plot of time-domain statistical features.

4.1.2 Frequency domain statistical analysis

As a second method, by applying Welch's PSD estimation to the current signal, analysis became possible in the frequency domain. After the statistical properties of the amplitudes in the frequency domain were extracted, classification algorithms were applied. As in the first method, statistical properties such as kurtosis, skewness, mean, RMS, standard deviation and median are extracted in the Knime for cases where the motor is healthy or faulty.

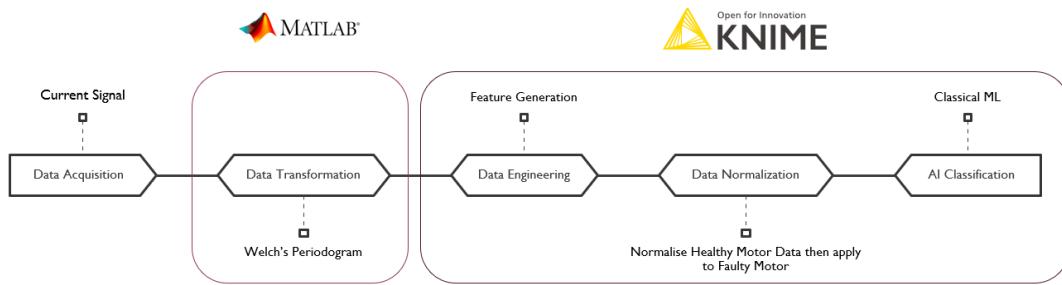


Figure 4.4 : Diagram of frequency domain statistical analysis method.

As can be seen from the t-SNE plot, a short-circuit fault between the stator windings and a bearing fault can be better differentiated from the healthy condition than the first method, while 1 bar-broken rotor fault is still relatively more difficult to distinguish. According to the classification performance results, it is seen that a high rate of success can still be achieved.

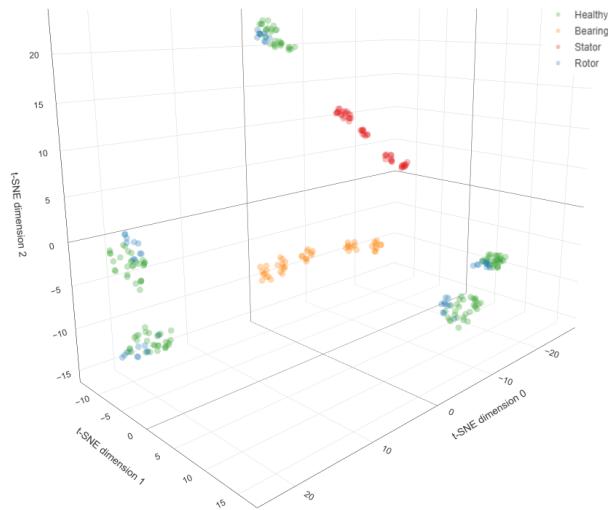


Figure 4.5 : t-SNE plot of frequency domain statistical features.

4.1.3 Statistical analysis on characteristic frequencies

As a third method, by applying Welch's PSD estimate to the current signal, analysis in the frequency domain became possible. Specific fault frequencies in the frequency domain were calculated, then the corresponding amplitudes were calculated. By applying statistical measures such as mean and standard deviation to the amplitudes, six features were created for each condition and classification algorithms were applied to these features. In this method, signal processing and feature extraction are computed in MATLAB, followed by classification in Knime.

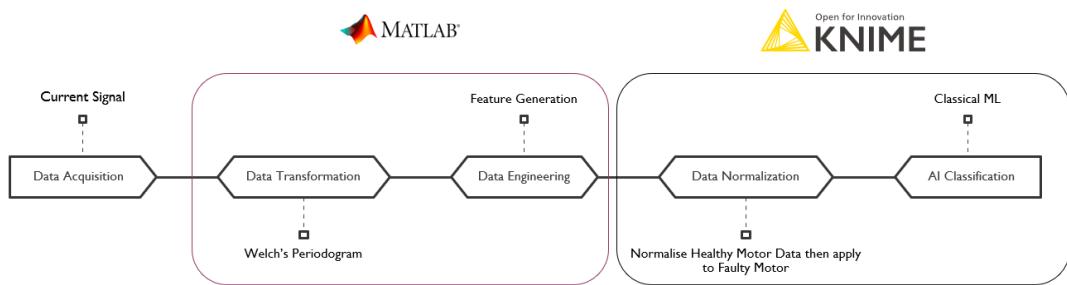


Figure 4.6 : Diagram of statistical analysis of characteristic frequencies method.

Condition monitoring and fault diagnosis researches for induction motors have been going on for many years and accordingly, there is a wide knowledge accumulation. Studies in the frequency domain have shown that fault conditions will have certain signs in the frequency spectrum. In this study, using the characteristic frequency equations for bearing, stator and rotor faults given in the literature review section (equations 2.3, 2.4 and 2.5), it has been calculated up to certain harmonics, and then the mean and standard deviation of the amplitudes corresponding to the frequencies obtained for each fault type are calculated and statistical features are formed.

As can be seen from the t-SNE plot, all faults can be better distinguished from the healthy state than the first and second method. According to the classification performance results, the third method outperforms the other two methods in almost every metric and every classifier.

4.1.4 Discussion

Due to their operating principles, machine learning methods are dependent on data engineering. Features revealed as a result of data engineering can better capture the fault characteristic. As can be seen from Table 4.1, statistical approaches generally

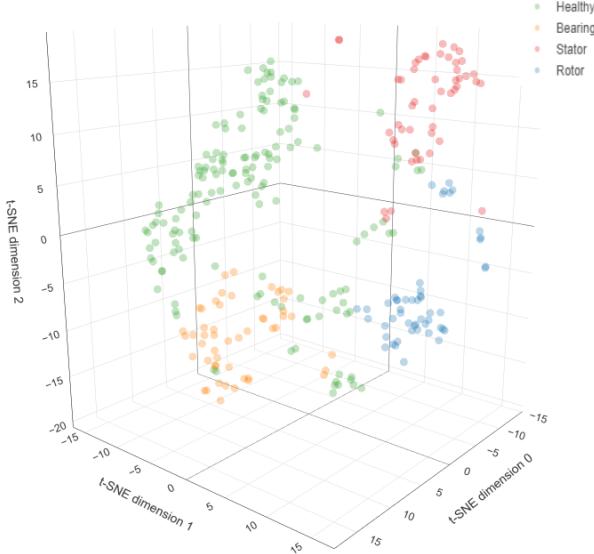


Figure 4.7 : t-SNE plot of characteristic frequencies statistics.

yield good results. Another advantage of statistical features is that they can be used in prognostic studies by observing their changes over time.

Although the statistical study (M1) in the time domain gives good results compared to its simplicity and computational load, it is more vulnerable to external influences. It is useful to approach it carefully, as it is prone to give false results under distorting influences.

PSD estimation with Welch's method is more resistant to disruptive effects due to its properties. While the higher computational load is a disadvantage, it has the potential to show effects that cannot be seen in the time domain. For this reason, it is also used especially in industrial standards. Considering the performance metrics, although the second method (M2) gives close results with the study in the time domain, it can be preferred because it is resistant to external effects.

The characteristic effects of faults on motor current have been known to academic and industrial researchers for a long time. The disadvantage of this method is that it gives good results only under nominal load and speed conditions. However, as revealed in this study, it can be said that it is both a high-performance and robust method by statistically examining the amplitudes at characteristic frequencies (M3) under different speed and load conditions. As Table 4.1 exhibits, it outperforms for each performance metric and each classifier compared to other methods.

Table 4.1 : Performance metrics for methods and classifiers with 5-fold cross-validation.

	AUC (Mean ± STD)			Cohen's Kappa (Mean ± STD)		
	M1	M2	M3	M1	M2	M3
MLP	0.976 (±0.017)	0.976 (±0.032)	0.988 (±0.019)	0.915 (±0.069)	0.833 (±0.064)	0.93 (±0.082)
SVM	0.954 (±0.029)	0.946 (±0.039)	0.988 (±0.016)	0.777 (±0.081)	0.809 (±0.044)	0.92 (±0.05)
Random Forest	0.995 (±0.004)	0.967 (±0.032)	0.997 (±0.003)	0.925 (±0.047)	0.897 (±0.072)	0.93 (±0.049)
XGBoost	0.986 (±0.014)	0.953 (±0.022)	0.992 (±0.007)	0.871 (±0.059)	0.847 (±0.079)	0.914 (±0.046)
Naive Bayes	0.948 (±0.019)	0.884 (±0.027)	0.989 (±0.004)	0.792 (±0.052)	0.662 (±0.146)	0.876 (±0.049)
kNN	0.982 (±0.015)	0.979 (±0.024)	0.991 (±0.014)	0.883 (±0.04)	0.894 (±0.035)	0.944 (±0.043)
Macro F-measure (Mean ± STD)						
	M1	M2	M3	M1	M2	M3
MLP	0.934 (±0.052)	0.882 (±0.049)	0.952 (±0.053)	0.943 (±0.047)	0.89 (±0.042)	0.953 (±0.055)
SVM	0.908 (±0.054)	0.856 (±0.029)	0.947 (±0.033)	0.86 (±0.048)	0.877 (±0.03)	0.947 (±0.033)
Random Forest	0.938 (±0.036)	0.922 (±0.061)	0.948 (±0.039)	0.95 (±0.031)	0.933 (±0.046)	0.953 (±0.032)
XGBoost	0.899 (±0.045)	0.893 (±0.052)	0.938 (±0.034)	0.913 (±0.04)	0.897 (±0.058)	0.943 (±0.03)
Naive Bayes	0.832 (±0.04)	0.893 (±0.132)	0.915 (±0.04)	0.867 (±0.031)	0.777 (±0.127)	0.917 (±0.033)
kNN	0.907 (±0.024)	0.925 (±0.029)	0.962 (±0.029)	0.923 (±0.025)	0.93 (±0.022)	0.963 (±0.027)
Accuracy (Mean ± STD)						
	M1	M2	M3	M1	M2	M3

4.2 Deep learning analysis

In the study, the Tensorflow library was used to develop deep learning models and the Keras interface, which is a software library working on Tensorflow for artificial neural networks, was used in the Knime Data Analytics tool.

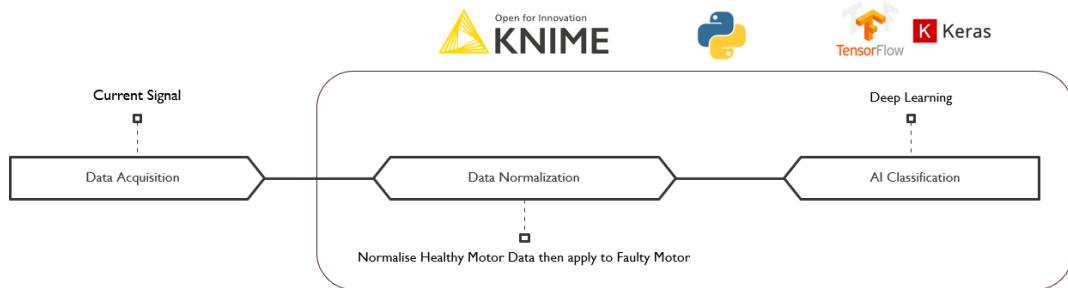


Figure 4.8 : Diagram of deep learning analysis.

Since machine learning methods require complex signal processing to extract the characteristics of the given data, the interest in the structures that will work end-to-end between the collected data and the decision output has increased in the industry. Deep learning methods can respond to this interest by not requiring any preprocessing. However, as a downside, large amounts of data are needed to train deep learning methods. In virtue of long-term data collection is not possible in the laboratory environment, the data set was divided into 0.2 seconds segments using the sliding window method with no-overlapping as in [86]. Original data length in time was 10 seconds, to ensure capturing fault characteristics equation 3.1 re-calculated and at least 40 times fault impacts captured. Each augmented segments were all assigned the same label as the original input sequence.

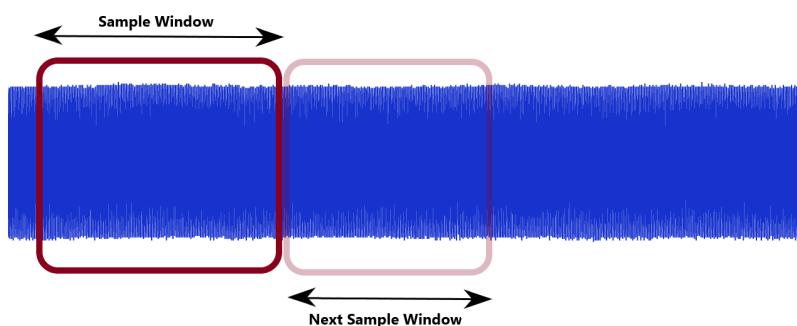


Figure 4.9 : Sliding window data augmentation.

The augmented dataset was used in the training of two different deep learning methods: Convolutional Neural Networks and LSTM Networks. Although CNN originated and mostly applied on 2-dimensional problems mainly in computer vision studies, 1-D CNN successfully applied on time series signals such as vibration, sound and current fault diagnosis of induction motors [82, 84, 96, 97].

An improved 1D-CNN structure for bearing fault diagnosis presented in [78], in this thesis study, structure improved further to the extended diagnosis of the stator winding short-circuit and broken rotor bar faults also.

Table 4.2 : 1D- CNN structure.

Type	Shape	Specific Setting
Input	(1000× 1)	5 kHz-0.2sec
Conv1D	(None, 1000, 128)	128@16x1, stride =1
BatchNormalization	(None, 1000, 128)	-
MaxPooling1D	(None, 500, 128)	pool size 2x1, stride =2
Dropout	(None, 500, 128)	0.25
Conv1D	(None, 500, 64)	64@8x1, stride =1
BatchNormalization	(None, 500, 64)	-
MaxPooling1D	(None, 250, 64)	pool size 2x1, stride =2
Dropout	(None, 250, 64)	0.25
Conv1D	(None, 250, 32)	32@4x1, stride =1
BatchNormalization	(None, 250, 32)	-
MaxPooling1D	(None, 125, 32)	pool size 2x1, stride =2
Dropout	(None, 125, 32)	0.1
Conv1D	(None, 125, 16)	16@4x1, stride =1
BatchNormalization	(None, 125, 16)	-
MaxPooling1D	(None, 62, 16)	pool size 2x1, stride =2
Dropout	(None, 62, 16)	0.1
Conv1D	(None, 62, 8)	8@4x1, stride =1
BatchNormalization	(None, 62, 8)	-
MaxPooling1D	(None, 31, 8)	pool size 2x1, stride =2
Dropout	(None, 31, 8)	0.25
Flatten	(None, 248)	
Dense	(None, 4)	

RNN is preferred in condition monitoring diagnostic applications because of its ability to work with sequential data. Since RNNs learn by backpropagation in time during a supervised learning process, problems with gradients during training can result in the increased error in time steps or distortion of prior knowledge [98]. In order to overcome this disadvantage, LSTMs have been developed to capture existing dependencies in time series data. Many studies have been carried out with

LSTM networks for fault detection and the effectiveness of the structure has been demonstrated [83, 85, 99, 100].

Table 4.3 : LSTM structure.

Type	Shape	Specific Setting
Input	(1000 × 1)	5 kHz-0.2sec units = 8
LSTM	(None, 1000, 8)	activation = 'tanh' recurrent activation = 'sigmoid'
BatchNormalization	(None, 1000, 8)	-
Dropout	(None, 1000, 8)	0.5
LSTM	(None, 1000, 16)	units = 16 activation = 'tanh' recurrent activation = 'sigmoid'
BatchNormalization	(None, 1000, 16)	-
Dropout	(None, 1000, 16)	0.4
LSTM	(None, 1000, 16)	units = 16 activation = 'tanh' recurrent activation = 'sigmoid'
BatchNormalization	(None, 1000, 16)	-
Dropout	(None, 1000, 16)	0.4
Flatten	(None, 16000)	
Dense	(None, 4)	

Detailed representations of 1D-CNN and LSTM structures are given in the appendix.

As given in Table 4.4, both deep learning methods provide impressive results in the diagnostics of faults. Comparing the two methods, 1D-CNN outperforms LSTM in all metrics, while LSTM is simpler and straightforward in structure. However, the time required to train both models is the same.

Table 4.4 : Performance metrics for deep learning architectures.

	AUC	Cohen's Kappa	Macro F-measure	Accuracy
1D-CNN	0.955	0.922	0.944	0.947
LSTM	0.936	0.87	0.904	0.912

4.3 Discussion

Machine learning methods can give better results than deep learning methods in relatively small datasets such as datasets obtained in experimental studies. As shown in Table 4.1, machine learning methods give successful results. It has been demonstrated that machine learning algorithms show high performance, especially when the features

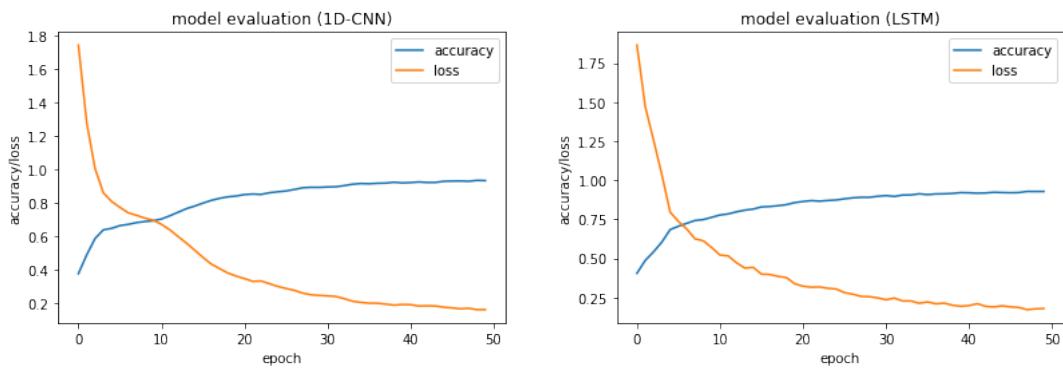


Figure 4.10a : Training performance of **Figure 4.10b :** Training performance of 1D-CNN model.

LSTM model.

made with statistical calculations at characteristic failure frequencies specified as the third method are used. However, it is obvious that it requires complex calculations and developer-based competence to achieve high performance.

As presented in Table 4.4, although deep learning methods are trained with much less data samples than machine learning methods, they gave very close results with machine learning methods trained with features extracted by the third method. In addition, it is another important issue that it provides such high performance without the need for any data preprocessing.

It can be said that the performance of deep learning methods will increase even more with the amount of data increasing with a high acceleration in industrial applications. Pre-trained deep learning models can also be useful in embedded systems with constrained resources. Especially, a deep learning-based condition monitoring and fault recognition application that will be integrated into motor drivers has great potential.

5. CONCLUSIONS AND RECOMMENDATIONS

The fact that induction motors have been in industrial applications for more than a century adds importance to condition monitoring studies. As the most common type of electric motor, it finds a place in a wide range from HVAC to the manufacturing and automotive industry.

In this study, first of all, the place and importance of asynchronous motors in the industrial field and the general operating principle are mentioned. Then, within the literature review, the failures that occurred due to the stresses they were exposed to during their operating life-cycle were explained. To avoid the negative consequences of failures with the least damage, effective maintenance methods are mentioned, and condition monitoring and fault recognition methods are explained as the backbone of efficient maintenance.

Within the scope of the study, bearing, stator winding short-circuit and broken rotor bar failures, which are the most common types of asynchronous motor failures, were created in laboratory conditions at WAT Motor facilities and their effects were examined. Studies on asynchronous motors fed directly and operating at nominal speed are common in the literature. In this study, tests were carried out at two different loadings, at different speeds from 30 Hz to 50 Hz, at 75% of the nominal and nominal load of the motor fed with the variable frequency drive.

With the obtained data, different methods and approaches have been examined for motor condition monitoring and fault detection. In the first of these, the statistical properties of the current signal were extracted in time-domain analysis. In the study conducted in the frequency domain, the statistical properties of the amplitudes obtained by estimating the Power Spectral Density with the Welch method were obtained in the same way. Finally, the amplitudes corresponding to the characteristic frequencies of the fault types in the frequency spectrum were calculated and the statistical properties of these amplitudes were extracted. With the features obtained by these three methods, machine learning classifiers were trained and motor fault diagnosis was carried out.

In short, a robust fault diagnosis method (M3), which is the statistical analysis of amplitudes of characteristic fault frequencies, proposed and compared with other machine learning pre-processing techniques. The presented method outperforms other techniques in all performance metrics. Also, two different deep learning structures, that are CNN and LSTM proposed in the diagnosis of bearing, stator and rotor faults in this study.

According to the findings, it is possible to get good results on fault diagnosis with statistical approaches. Considering the industrial application conditions, electric motors are heavily exposed to disruptive effects. PSD estimation with the Welch method gives very robust results against disruptive effects due to its structure. The amplitude-based statistical approach obtained from the PSD frequency spectrum with the fault characteristic formulas outperformed the other two methods in all metrics and with high accuracy and precision.

Deep learning studies and applications, which have been increasing in recent years, also find a place for themselves in the diagnostics of motor faults. As an alternative to feature extraction and signal processing problems that require time and expert knowledge in machine learning methods, deep learning methods can offer an end-to-end solution. As an alternative to data engineering and machine learning methods mentioned within the scope of the thesis, fault diagnosis was made with two deep learning methods, convolutional and recurrent neural networks. Deep learning methods show high performance without the need for any preprocessing, but as a downside, they need high processing power and a large dataset for training the model.

Although deep learning methods need large datasets during their training, they can work with less dimensional data than machine learning methods, as shown in the thesis. In this respect, a trained deep learning algorithm to be integrated into the VFD can work with less resource need via its feedforward structure.

In this study, experimental conditions were set for steady-state and analyzes were made for this state. For future studies, it will be more useful to analyze data under non-stationary state or to work on real state data from motors in the field. As another issue, the classification studies carried out can be integrated into an embedded system. In particular, the machine learning models, which are trained with statistical features

applied to characteristic frequencies as in this thesis, and deep learning methods will be useful in future VFD applications.

As a conclusion within Industry 4.0, some concepts become even more important. While the first two of them are facilitating access to data and increasing power on data processing, the third one can be said as efficiency. In the case of asynchronous motors, these three concepts can be met with VFDs. While the efficiency of the system increases in motors fed with VFD, the current signal required for motor control can be used for condition monitoring studies without any additional expense. It has the potential to facilitate data access by transferring the collected data to data centres with the Internet of Things infrastructures.

REFERENCES

- [1] **Albrecht, P., Appiarius, J., Cornell, E., Houghtaling, D., McCoy, R., Owen, E. and Sharma, D.** (1987). Assessment of the reliability of motors in utility applications, *IEEE transactions on energy conversion*, (3), 396–406.
- [2] **EN, B.** (2017). 13306:2017: Maintenance—Maintenance terminology, *BSI Standards Publication*.
- [3] **Tinga, T. and Loendersloot, R.**, (2019). Physical model-based prognostics and health monitoring to enable predictive maintenance, Predictive Maintenance in Dynamic Systems, Springer, pp.313–353.
- [4] **Karmakar, S., Chatopadhyay, S., Mitra, M. and Sengupta, S.** *Induction motor fault diagnosis*, volume 25, Springer.
- [5] **Jwo, D.J., Chang, W.Y. and Wu, I.H.** (2021). Windowing Techniques, the Welch Method for Improvement of Power Spectrum Estimation, *CMC-COMPUTERS MATERIALS & CONTINUA*, 67(3), 3983–4003.
- [6] **Solomon Jr, O.** (1991). PSD computations using Welch's method, *NASA STI/Recon Technical Report N*, 92, 23584.
- [7] **Waide, P. and Brunner, C.U.** (2011). Energy-efficiency policy opportunities for electric motor-driven systems.
- [8] **Kulterer, K., Werle, R., Lackner, P., Brunner, C. and Ellis, M.** (2014). Policy Guidelines for Electric Motor Systems—Part 2: Toolkit for Policy Makers, *4E Electric Motor Systems EMSA, 4E Energy Efficient Enduse Equipment, International Energy Agency*.
- [9] **Fleiter, T. and Eichhammer, W.** (2012). Energy efficiency in electric motor systems : Technology , saving potentials and policy options for developing countries.
- [10] **Mikami, H., Ide, K., Shimizu, Y., Senoo, M. and Seki, H.** (2011). Historical evolution of motor technology, *Hitachi Review*, 60(1), 39.
- [11] (2014). Improving Motor and Drive System Performance – A Sourcebook for Industry, <https://www.osti.gov/biblio/1220836>.
- [12] (1985). Report of Large Motor Reliability Survey of Industrial and Commercial Installations, Part I, *IEEE Transactions on Industry Applications*, IA-21(4), 853–864.

- [13] **Hovstadius, G. and Bolles, S.** (2016). Manual for Industrial Pump Systems Assessment and Optimization, **Technical Report**, United Nations Industrial Development Organization.
- [14] **Oliver, J.** (1992). Electric motor predictive and preventive maintenance guide, **Technical Report**, Electric Power Research Inst.
- [15] **Faiz, J., Joksimović, G. and Ghorbanian, V.** (2017). *Fault diagnosis of induction motors*, Institution of Engineering & Technology.
- [16] **Motors, W.E.** (2018). Induction motors fed by PWM frequency inverters, **Technical Report**, <https://www.weg.net/institutional/US/en/search/downloadcenter>.
- [17] **Drives, D.** (2019). Facts Worth Knowing about AC Drives, **Technical Report**, <https://danfoss.ipapercms.dk/Drives/DD/Global/SalesPromotion/FWK>.
- [18] **Bose, B.** (2002). *Modern Power Electronics and AC Drives*, Eastern Economy Edition, Prentice Hall PTR.
- [19] **Lughofer, E. and Sayed-Mouchaweh, M.**, (2019). Prologue: Predictive maintenance in dynamic systems, Predictive Maintenance in Dynamic Systems, Springer, pp.1–23.
- [20] **Ruiz-Sarmiento, J.R., Monroy, J., Moreno, F.A., Galindo, C., Bonelo, J.M. and Gonzalez-Jimenez, J.** (2020). A predictive model for the maintenance of industrial machinery in the context of industry 4.0, *Engineering Applications of Artificial Intelligence*, 87, 103289.
- [21] **Ahmad, R. and Kamaruddin, S.** (2012). An overview of time-based and condition-based maintenance in industrial application, *Computers & Industrial Engineering*, 63(1), 135–149.
- [22] **Bonnett, A.H.** (2010). Root cause failure analysis for AC Induction Motors in the petroleum and chemical industry, *2010 Record of Conference Papers Industry Applications Society 57th Annual Petroleum and Chemical Industry Conference (PCIC)*, pp.1–13.
- [23] **Albrecht, P., Appiarius, J., McCoy, R., Owen, E. and Sharma, D.** (1986). Assessment of the reliability of motors in utility applications-Updated, *IEEE Transactions on Energy conversion*, (1), 39–46.
- [24] **Thorsen, O.V. and Dalva, M.** (1995). A survey of faults on induction motors in offshore oil industry, petrochemical industry, gas terminals, and oil refineries, *IEEE transactions on industry applications*, 31(5), 1186–1196.
- [25] **Bonnett, A.H. and Yung, C.** (2008). Increased efficiency versus increased reliability, *IEEE Industry Applications Magazine*, 14(1), 29–36.
- [26] **Trigeassou, J.C.** (2013). *Electrical machines diagnosis*, John Wiley & Sons.

- [27] **Zhang, P., Du, Y., Habetler, T.G. and Lu, B.** (2010). A survey of condition monitoring and protection methods for medium-voltage induction motors, *IEEE Transactions on Industry Applications*, 47(1), 34–46.
- [28] **Bonnet, A.H.** (2012). The Cause and Analysis of Bearing and Shaft Failures in Electric Motors, **Technical Manual 27**, EASA Technology & Education Consultant.
- [29] (2017). Bearing damage and failure analysis, **Technical Manual PUB BU/I3 14219/2 EN**, SKF Group.
- [30] **Schoen, R.R., Habetler, T.G., Kamran, F. and Bartfield, R.** (1995). Motor bearing damage detection using stator current monitoring, *IEEE transactions on industry applications*, 31(6), 1274–1279.
- [31] **EN, B.** (2013). 20958:2013: Condition monitoring and diagnostics of machine systems—Electrical signature analysis of three-phase induction motors, *BSI Standards Publication*.
- [32] **Siddique, A., Yadava, G. and Singh, B.** (2005). A review of stator fault monitoring techniques of induction motors, *IEEE Transactions on Energy Conversion*, 20(1), 106–114.
- [33] **Lipo, T.A.** (2017). *Introduction to AC Machine Design*, John Wiley & Sons, Ltd.
- [34] **Bonnett, A.H.** (1999). Root cause AC motor failure analysis, *Industry Applications Society 46th Annual Petroleum and Chemical Technical Conference (Cat. No. 99CH37000)*, IEEE, pp.85–97.
- [35] **Penman, J., Sedding, H., Lloyd, B. and Fink, W.** (1994). Detection and location of interturn short circuits in the stator windings of operating motors, *IEEE Transactions on Energy Conversion*, 9(4), 652–658.
- [36] **İmeryüz, M.** (2009). Asenkron Makinede Kafes Kırılmasının İncelenmesi, *Ph.D. Dissertation*, Fen Bilimleri Enstitüsü.
- [37] **Filippetti, F., Bellini, A. and Capolino, G.A.** (2013). Condition monitoring and diagnosis of rotor faults in induction machines: State of art and future perspectives, *2013 IEEE Workshop on Electrical Machines Design, Control and Diagnosis (WEMDCD)*, IEEE, pp.196–209.
- [38] **Mistry, R., Finley, W.R., Hashish, E. and Kreitzer, S.** (2016). Rotating machines—Pros and cons of monitoring devices, *2016 Petroleum and Chemical Industry Technical Conference (PCIC)*, IEEE, pp.1–12.
- [39] **Society, I.P.E.** (2017). IEEE standard test procedure for polyphase induction motors and generators, *ANSI, IEEE Standard 112*.
- [40] **Kumar, S., Mukherjee, D., Guchhait, P.K., Banerjee, R., Srivastava, A.K., Vishwakarma, D. and Saket, R.** (2019). A comprehensive review of condition based prognostic maintenance (CBPM) for induction motor, *Ieee Access*, 7, 90690–90704.

- [41] **Thorsen, O. and Dalva, M.** (1998). Methods of condition monitoring and fault diagnosis for induction motors, *European transactions on electrical power*, 8(5), 383–395.
- [42] **Randall, R.B.** (2021). *Vibration-based condition monitoring: industrial, automotive and aerospace applications*, John Wiley & Sons.
- [43] **EN, I.** (2002). 13373-1:2002: Condition monitoring and diagnostics of machines — Vibration condition monitoring — Part 1: General procedures, *International Organization for Standardization*.
- [44] **Thomson, W.T. and Fenger, M.** (2001). Current signature analysis to detect induction motor faults, *IEEE Industry Applications Magazine*, 7(4), 26–34.
- [45] **Gritli, Y., Bellini, A., Rossi, C., Casadei, D., Filippetti, F. and Capolino, G.** (2017). Condition monitoring of mechanical faults in induction machines from electrical signatures: Review of different techniques, *2017 IEEE 11th international symposium on diagnostics for electrical machines, power electronics and drives (SDEMPED)*, IEEE, pp.77–84.
- [46] **Corne, B., Knockaert, J. and Desmet, J.** (2017). Misalignment and unbalance fault severity estimation using stator current measurements, *2017 IEEE 11th International Symposium on Diagnostics for Electrical Machines, Power Electronics and Drives (SDEMPED)*, IEEE, pp.247–253.
- [47] **Shaeboub, A.** (2018). The Monitoring of Induction Machines Using Electrical Signals from the Variable Speed Drive, *Ph.D. Dissertation*, University of Huddersfield.
- [48] **Cernuda, C.**, (2019). On the relevance of preprocessing in predictive maintenance for dynamic systems, *Predictive Maintenance in Dynamic Systems*, Springer, pp.53–93.
- [49] **Bonaldi, E.L., de Oliveira, L.E.d.L., da Silva, J.G.B., Lambert-Torresm, G. and da Silva, L.E.B.**, (2012). Predictive maintenance by electrical signature analysis to induction motors, *Induction Motors-Modelling and Control*, IntechOpen.
- [50] **Medjaher, K., Camci, F. and Zerhouni, N.** (2012). Feature extraction and evaluation for Health Assessment and Failure prognostics., *Proceedings of First European Conference of the Prognostics and Health Management Society, PHM-E'12.*, <https://www.phmsociety.org/events/conference/europhm/12/proceedings>, Anibal Bregon, Abhinav Saxena, pp.111–116.
- [51] **Soualhi, M., Nguyen, K.T., Soualhi, A., Medjaher, K. and Hemsas, K.E.** (2019). Health monitoring of bearing and gear faults by using a new health indicator extracted from current signals, *Measurement*, 141, 37–51.

- [52] **Shukla, S., Yadav, R., Sharma, J. and Khare, S.** (2015). Analysis of statistical features for fault detection in ball bearing, *2015 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC)*, IEEE, pp.1–7.
- [53] **Croarkin, C. and Tobias, P.** (2012). e-Handbook of Statistical Methods, *NIST/SEMATECH*.
- [54] **Sait, A.S. and Sharaf-Eldeen, Y.I.** (2011). A review of gearbox condition monitoring based on vibration analysis techniques diagnostics and prognostics, *Rotating Machinery, Structural Health Monitoring, Shock and Vibration, Volume 5*, 307–324.
- [55] **Allen, R.L. and Mills, D.** (2004). *Signal analysis: time, frequency, scale, and structure*, John Wiley & Sons.
- [56] **Ahmed, H. and Nandi, A.K.** (2019). *Condition Monitoring with Vibration Signals*, Wiley Online Library.
- [57] **Orfanidis, S.J.** (1995). *Introduction to signal processing*, Prentice-Hall, Inc.
- [58] **Hayes, M.H.** (2009). *Statistical digital signal processing and modeling*, John Wiley & Sons.
- [59] **CusidÓCusido, J., Romeral, L., Ortega, J.A., Rosero, J.A. and Espinosa, A.G.** (2008). Fault detection in induction machines using power spectral density in wavelet decomposition, *IEEE Transactions on Industrial Electronics*, 55(2), 633–643.
- [60] **Irvine, T.** (2002). *An introduction to shock & vibration response spectra*.
- [61] **Zerdani, S., El Hafyani, M.L. and Zouggar, S.** (2020). Inter-Turn Stator Winding fault Diagnosis for Permanent Magnet Synchronous Motor based Power Spectral Density Estimators, *2020 International Conference on Smart Grid and Clean Energy Technologies (ICSGCE)*, IEEE, pp.137–142.
- [62] **Heydarzadeh, M., Madani, N. and Nourani, M.** (2016). Gearbox fault diagnosis using power spectral analysis, *2016 IEEE International Workshop on Signal Processing Systems (SiPS)*, IEEE, pp.242–247.
- [63] **Stoica, P., Moses, R.L. et al.** (2005). *Spectral analysis of signals*, Pearson Prentice Hall Upper Saddle River, NJ.
- [64] **Schmid, H.** (2012). How to use the FFT and Matlab's pwelch function for signal and noise simulations and measurements, *FHNW/IME*, 2–13.
- [65] **Al Ahmar, E., Choqueuse, V., Benbouzid, M., Amirat, Y., El Assad, J., Karam, R. and Farah, S.** (2010). Advanced signal processing techniques for fault detection and diagnosis in a wind turbine induction generator drive train: A comparative study, *2010 IEEE Energy Conversion Congress and Exposition*, IEEE, pp.3576–3581.

- [66] **Jin, Z., Han, Q., Zhang, K. and Zhang, Y.** (2020). An intelligent fault diagnosis method of rolling bearings based on Welch power spectrum transformation with radial basis function neural network, *Journal of Vibration and Control*, 26(9-10), 629–642.
- [67] **Ayhan, B., Chow, M.Y., Trussell, H.J. and Song, M.H.** (2003). A case study on the comparison of non-parametric spectrum methods for broken rotor bar fault detection, *IECON'03. 29th Annual Conference of the IEEE Industrial Electronics Society (IEEE Cat. No. 03CH37468)*, volume 3, IEEE, pp.2835–2840.
- [68] **Friedman, N., Geiger, D. and Goldszmidt, M.** (1997). Bayesian network classifiers, *Machine learning*, 29(2), 131–163.
- [69] **Lowd, D. and Domingos, P.** (2005). Naive Bayes models for probability estimation, *Proceedings of the 22nd international conference on Machine learning*, pp.529–536.
- [70] **Martin-Diaz, I., Morinigo-Sotelo, D., Duque-Perez, O. and Romero-Troncoso, R.J.** (2018). An experimental comparative evaluation of machine learning techniques for motor fault diagnosis under various operating conditions, *IEEE Transactions on Industry Applications*, 54(3), 2215–2224.
- [71] **Shalev-Shwartz, S. and Ben-David, S.** (2014). *Understanding machine learning: From theory to algorithms*, Cambridge university press.
- [72] **Richman, J.S.** (2011). Multivariate neighborhood sample entropy: a method for data reduction and prediction of complex data, *Methods in enzymology*, 487, 397–408.
- [73] **Liu, R., Yang, B., Zio, E. and Chen, X.** (2018). Artificial intelligence for fault diagnosis of rotating machinery: A review, *Mechanical Systems and Signal Processing*, 108, 33–47.
- [74] **Haykin, S.** (2010). *Neural networks and learning machines*, 3/E, Pearson Education India.
- [75] **Alpaydin, E.** (2020). *Introduction to machine learning*, MIT press.
- [76] **Sewell, M.** (2008). Ensemble learning, *RN*, 11(02), 1–34.
- [77] **Breiman, L.** (2001). Random forests, *Machine learning*, 45(1), 5–32.
- [78] **Chen, C.C., Liu, Z., Yang, G., Wu, C.C. and Ye, Q.** (2021). An Improved Fault Diagnosis Using 1D-Convolutional Neural Network Model, *Electronics*, 10(1), 59.
- [79] **Company, W.M.**, (2021), <https://www.wat.com.tr/UPLOAD/CONTENT/dokumanlar/brosurlerimiz/Cast%20Iron%20Motor%20Series.pdf>.

- [80] **Zhang, S., Zhang, S., Wang, B. and Habetler, T.G.** (2020). Deep learning algorithms for bearing fault Diagnosticsx—A comprehensive review, *IEEE Access*, 8, 29857–29881.
- [81] **Zhao, R., Yan, R., Chen, Z., Mao, K., Wang, P. and Gao, R.X.** (2019). Deep learning and its applications to machine health monitoring, *Mechanical Systems and Signal Processing*, 115, 213–237.
- [82] **Skowron, M., Orlowska-Kowalska, T., Wolkiewicz, M. and Kowalski, C.T.** (2020). Convolutional neural network-based stator current data-driven incipient stator fault diagnosis of inverter-fed induction motor, *Energies*, 13(6), 1475.
- [83] **Khan, S. and Yairi, T.** (2018). A review on the application of deep learning in system health management, *Mechanical Systems and Signal Processing*, 107, 241–265.
- [84] **Eren, L., Ince, T. and Kiranyaz, S.** (2019). A generic intelligent bearing fault diagnosis system using compact adaptive 1D CNN classifier, *Journal of Signal Processing Systems*, 91(2), 179–189.
- [85] **Sabir, R., Rosato, D., Hartmann, S. and Guehmann, C.** (2019). Lstm based bearing fault diagnosis of electrical machines using motor current signal, *2019 18th IEEE International Conference On Machine Learning And Applications (ICMLA)*, IEEE, pp.613–618.
- [86] **Shenfield, A. and Howarth, M.** (2020). A novel deep learning model for the detection and identification of rolling element-bearing faults, *Sensors*, 20(18), 5112.
- [87] **Canbek, G., Sagiroglu, S., Temizel, T.T. and Baykal, N.** (2017). Binary classification performance measures/metrics: A comprehensive visualized roadmap to gain new insights, *2017 International Conference on Computer Science and Engineering (UBMK)*, IEEE, pp.821–826.
- [88] **Seliya, N., Khoshgoftaar, T.M. and Van Hulse, J.** (2009). A study on the relationships of classifier performance metrics, *2009 21st IEEE international conference on tools with artificial intelligence*, IEEE, pp.59–66.
- [89] **Grandini, M., Bagli, E. and Visani, G.** (2020). Metrics for Multi-Class Classification: an Overview, *arXiv preprint arXiv:2008.05756*.
- [90] **Han, T., Liu, C., Wu, R. and Jiang, D.** (2021). Deep transfer learning with limited data for machinery fault diagnosis, *Applied Soft Computing*, 103, 107150.
- [91] **He, H. and Garcia, E.A.** (2009). Learning from imbalanced data, *IEEE Transactions on knowledge and data engineering*, 21(9), 1263–1284.
- [92] **Janssens, O., Slavkovikj, V., Vervisch, B., Stockman, K., Loccufier, M., Verstockt, S., Van de Walle, R. and Van Hoecke, S.**

- (2016). Convolutional neural network based fault detection for rotating machinery, *Journal of Sound and Vibration*, 377, 331–345.
- [93] **Widmann, M. and Roccato, A.** (2021). *From Modeling to Model Evaluation*, Knime Press.
- [94] **Fawcett, T.** (2004). ROC graphs: Notes and practical considerations for researchers, *Machine learning*, 31(1), 1–38.
- [95] **Van der Maaten, L. and Hinton, G.** (2008). Visualizing data using t-SNE., *Journal of machine learning research*, 9(11).
- [96] **Eren, L.** (2017). Bearing fault detection by one-dimensional convolutional neural networks, *Mathematical Problems in Engineering*, 2017.
- [97] **Qiao, M., Yan, S., Tang, X. and Xu, C.** (2020). Deep convolutional and LSTM recurrent neural networks for rolling bearing fault diagnosis under strong noises and variable loads, *IEEE Access*, 8, 66257–66269.
- [98] **Enshaei, N. and Naderkhani, F.** (2019). Application of deep learning for fault diagnostic in induction Machine's bearings, *2019 IEEE International Conference on Prognostics and Health Management (ICPHM)*, IEEE, pp.1–7.
- [99] **Bai, M., Liu, J., Ma, Y., Zhao, X., Long, Z. and Yu, D.** (2021). Long short-term memory network-based normal pattern group for fault detection of three-shaft marine gas turbine, *Energies*, 14(1), 13.
- [100] **He, D., He, J., Liu, J., Yang, J., Yan, Q. and Yang, Y.** (2021). An FPGA-Based LSTM Acceleration Engine for Deep Learning Frameworks, *Electronics*, 10(6), 681.

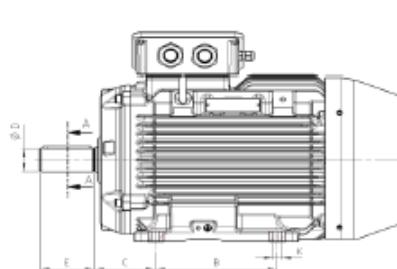
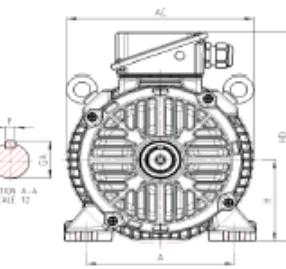
APPENDICES

APPENDIX A.1 : Induction motor data sheet

APPENDIX B.1 : Parameters of machine learning algorithms

APPENDIX C.1 : Diagrams of deep learning algorithms

APPENDIX A.1

	WAT Motor Sanayi ve Ticaret A.Ş.																								
Three Phase Squirrel Cage Induction Motor																									
Motor type P3HG160M4C40-KI																									
Electrical Data																									
RATED VALUES				STARTING VALUES				BREAKDOWN		EFFICIENCY			EFFICIENCY CLASS												
VOLTAGE	POWER	SPEED	CURRENT	TORQUE	Cos φ	4/4	I _s / I _n	M _A / M _N	Y D	M _K / M _N	Nm	4/4	3/4	2/4	IEC 60034-30-1										
V	kW	Hz	rpm	A	Nm	4/4	Y D	M _A / M _N	Y D	M _K / M _N	Nm	4/4	3/4	2/4	IE										
400/690	11	50	1475	22,0/12,7	71,30	0,79	2,4	7,3	0,9	2,7	3,1	91,4	91,6	90,7	IE3										
460/795	12,7	60	1775	21,65/7,2	68,23	0,8	2,4	7,06	0,9	2,5	3,0	-	-	-											
Mechanical Data																									
Bearing DE	6309-2Z C3				Color				RAL 3009																
Bearing NDE	6309-2Z C3				Motor protection				PTC																
Lubricants	Standart				Method of cooling				IC411 (TEPC)																
Condensation drain hole	Optional				Terminal box position				Top																
External earthing terminal	Yes				Material of terminal box				Cast Iron																
Vibration class	A				Terminal plate thread				M6																
Insulation Class	F				Cable gland				2xM32G+1xM12G																
Duty type	S1				Clamping range				18-25 mm																
Direction of rotation	CW/CCW				Mounting type				IM 1001 (B3)																
Housing material	Cast Iron																								
Type of balancing	H																								
Environmental Data							Others																		
Ambient Temperature	-20 / +40 °C				Type of protection				IP 56																
Altitude above sea level	1000 m				Sound Pressure Level 50Hz dBA				62																
					Moment of Inertia (kg.mm ²)				0,07542025																
					Standard				IEC 60034-1																
Dimensions																									
																									
Main Dimensions			Foot Mounted Motors					Shaft			Flange Mounted Motors														
AC	L	B	A	H	HD	K	C	D	E	GA	F	P	N	M	S										
328	626	210	254	160	387	14,5	108	42	110	45	12	-	-	-	-										

APPENDIX B.1

Table B.1 : Naive Bayes algorithm parameters.

Parameter	Value
Default Probability	0.005
Minimum Standard Deviation	0.001
Threshold Standard Deviation	0.001
Maximum number of unique nominal values per attribute	25

Table B.2 : Multi-Layer Perceptron algorithm parameters.

Parameter	Value
Number of iterations	100
Number of hidden layers	2
Number of neurons per layers	20

Table B.3 : Random Forest algorithm parameters.

Parameter	Value
Split criterion	Information Gain
Limit number of levels	8
Number of minimum node size	2
Number of models	1000

Table B.4 : k-Nearest Neighbors algorithm parameter.

Parameter	Value
Number of neighbor to consider	5

Table B.5 : Support Vector Machines algorithm parameters.

Parameter	Value
Type	C-SVC
Kernel	Linear
Cost	0.9
Epsilon	0.0001

Table B.6 : Extreme Gradient Boosting algorithm parameters.

Parameter	Value
Boosting Rounds	1000
Objective	Softprob
Booster	Tree
Eta	0.3
Gamma	0
Max. depth	6
Min. child weight	1
Max.. delta step	0
Subsampling rate	1
Column sampling rate by tree	1
Column sampling rate by level	1
Lambda	1
Alpha	1
Tree method	Auto
Sketch epsilon	0.03
Scale by positive weight	1
Ground pooling	DeptWise
Max. number of leaves	0
Max. number of bins	256

APPENDIX C.1

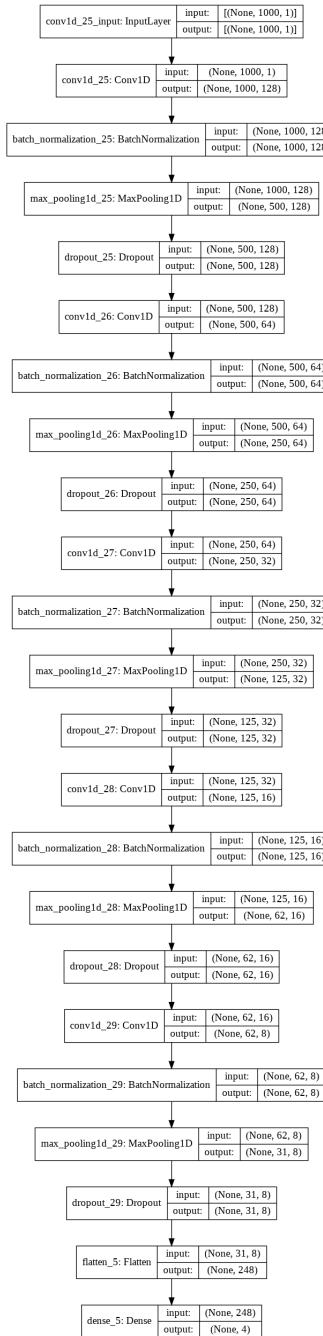


Figure C.1 : 1D-CNN model diagram.

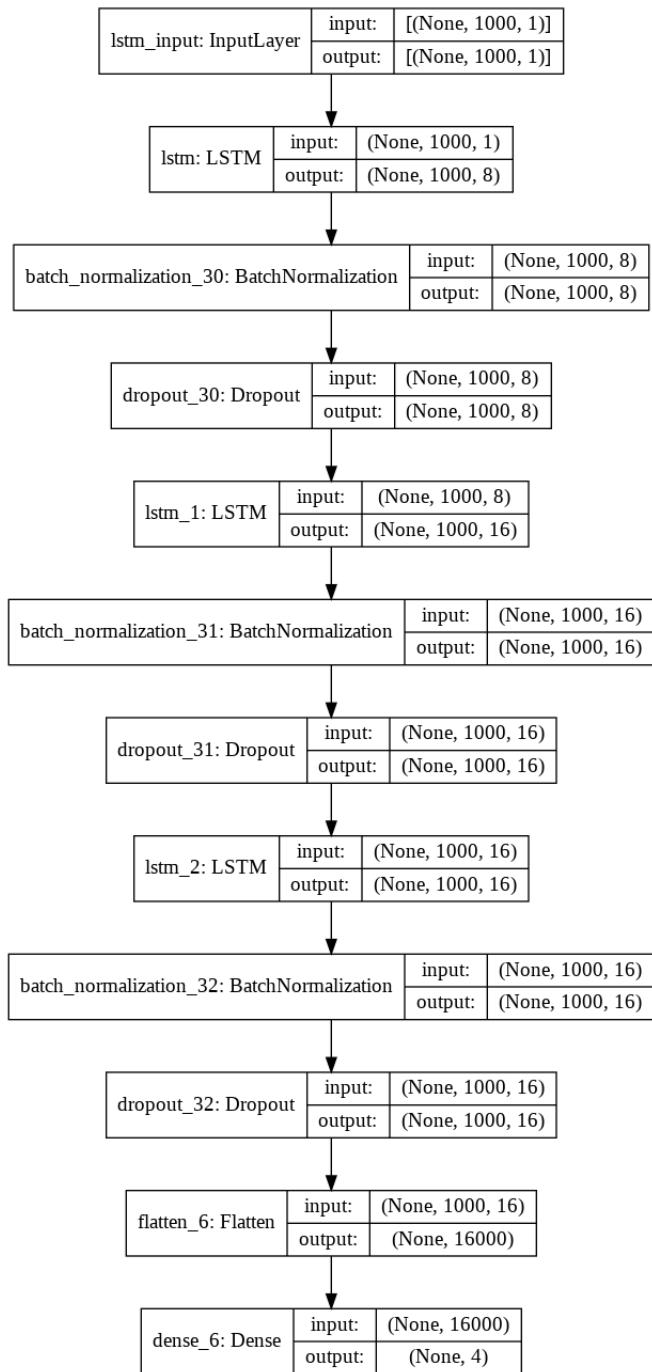
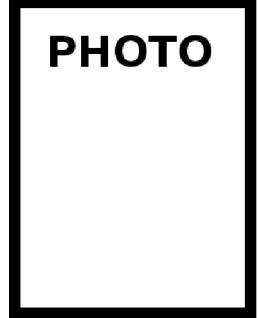


Figure C.2 : LSTM model diagram.

CURRICULUM VITAE

PHOTO



Name Surname : Alper Senem

Place and Date of Birth : Ankara - 15.06.1994

E-Mail : senemal@itu.edu.tr

EDUCATION :

- **B.Sc.** : 2018, Istanbul Technical University, Faculty of Mechanical Engineering, Department of Mechanical Engineering

PROFESSIONAL EXPERIENCE AND REWARDS:

- 2019-2020 Arcelik Global / WAT Motor R&D Department
Research and Development Engineer
- 2020-2021 Arcelik Global / Central R&D Department
Research and Development Engineer

PUBLICATIONS, PRESENTATIONS AND PATENTS ON THE THESIS:

- **Ertuğrul Ş, Senem A.** 2019. Condition Monitoring and Fault Diagnosis of Induction Motors: A Survey *International Conference on Artificial Intelligence towards Industry 4.0 (ICAI4.0)*, November, 2019 Iskenderun, Hatay / Turkey.

ALPER SENEM

DATA-DRIVEN CONDITION MONITORING AND FAULT
DIAGNOSIS OF VFD-FED INDUCTION MOTORS

2021