Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies

Bc. Miroslav Hájek

Machinery vibrodiagnostics with the industrial internet of things

Progress report in Master's thesis project I

Thesis Supervisor: Ing. Marcel Baláž, PhD.

June 2023

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies

Bc. Miroslav Hájek

Machinery vibrodiagnostics with the industrial internet of things

Progress report in Master's thesis project I

Study programme:	Intelligent Software Systems
Study field:	Informatics
Training workplace:	Institute of Computer Engineering and Applied Informatics
Thesis supervisor:	Ing. Marcel Baláž, PhD.
Departmental advisor:	Ing. Jakub Findura
Consultant:	Ing. Lukáš Doubravský

June 2023

SLOVENSKÁ TECHNICKÁ UNIVERZITA V BRATISLAVE
FAKULTA INFORMATIKY A INFORMAČNÝCH TECHNOLÓGIÍ

Návrh zadania diplomovej práce

Finálna verzia do diplomovej práce [1]

Študent:

Meno, priezvisko, tituly:	Miroslav Hájek, Bc.
Študijný program:	Inteligentné softvérové systémy
Kontakt:	xhajekm@stuba.sk

Výskumník:

Meno, priezvisko, tituly:	Marcel Baláž, Ing. PhD.

Projekt:

Názov:	Vibrodiagnostika strojov s priemyselným internetom vecí
Názov v angličtine:	Machinery vibrodiagnostics with the industrial internet of things
Miesto vypracovania:	Ústav počítačového inžinierstva a aplikovanej informatiky, FIIT STU, Bratislava
Oblasť problematiky:	Internet vecí, Spracovanie signálov, Feature engineering

Text návrhu zadania[2]

Monitorovanie prevádzkového stavu rotačných strojov za účelom včasného odhalenia poškodení je dôležité pre plynulý priebeh priemyselných procesov bez náhleho zlyhania kľúčového technického vybavenia. Nadmerné vibrácie alebo graduálna či náhla zmena ich charakteru sú spoľahlivými indikátormi opotrebenia dielcov. V mnohých prípadoch bývajú zavedené iba pravidelné pôchodzkové merania s následným vyhodnotením časových a frekvenčných priebehov kvalifikovaným personálom. Kontinuálna diagnostika a prediktívna údržba rozširujúca sa so zariadeniami IIoT spôsobuje enormný nárast objemu zaznamenaných dát. Sledovanie výchyliek operátorom a manuálna identifikácia súčiastok vyžadujúcich údržbu v celom závode sa tak stáva prakticky nerealizovateľná.

Preskúmajte spôsoby zisťovania bežných poškodení strojov z vibračných signálov a analyzujte algoritmy na redukciu množstva posielaných dát zo senzorov vzhľadom na osobitosti aplikačnej domény. Navrhnite reprezentáciu údajov na základe typických čŕt signálu, ktorá zníži výpočtové nároky na zvyšok komunikačného reťazca. Zvolený spôsob predspracovania má zároveň umožniť diagnostiku poškodení zvoleného stroja. Implementuje vaše riešenie s ohľadom na možné nasadenie na prostriedkami limitovanú senzorovú jednotku. Následne posúďte efektívnosť, porovnajte dosiahnuté presnosti diagnostiky a verifikujte voči zaužívaným postupom.

---

[1] Vytlačiť obojstranne na jeden list papiera

[2] 150-200 slov (1200-1700 znakov), ktoré opisujú výskumný problém v kontexte súčasného stavu vrátane motivácie a smerov riešenia

Literatúra[3]

- NANDI, Asoke Kumar; AHMED, Hosameldin. Condition monitoring with vibration signals: compressive sampling and learning algorithms for rotating machines. Hoboken, NJ, USA: Wiley-IEEE Press, 2019. ISBN 978-1-119-54462-3.
- YU, Gang. A Concentrated Time–Frequency Analysis Tool for Bearing Fault Diagnosis. IEEE Transactions on Instrumentation and Measurement. 2020, vol. 69, no. 2, pp. 371–381. ISSN 1557-9662. DOI: 10.1109/TIM.2019.2901514. Conference Name: IEEE Transactions on Instrumentation and Measurement.

Vyššie je uvedený návrh diplomového projektu, ktorý vypracoval(a) Bc. Miroslav Hájek, konzultoval(a) a osvojil(a) si ho Ing. Marcel Baláž, PhD. a súhlasí, že bude takýto projekt viesť v prípade, že bude pridelený tomuto študentovi.

V Bratislave dňa 22.2.2023

Podpis študenta		Podpis výskumníka

Vyjadrenie garanta predmetov Diplomový projekt I, II, III

Návrh zadania schválený: áno / nie[4]

Dňa: ....

Podpis garanta predmetov

[3] 2 vedecké zdroje, každý v samostatnej rubrike a s údajmi zodpovedajúcimi bibliografickým odkazom podľa normy STN ISO 690, ktoré sa viažu k téme zadania a preukazujú výskumnú povahu problému a jeho aktuálnosť (uveďte všetky potrebné údaje na identifikáciu zdroja, pričom uprednostnite vedecké príspevky v časopisoch a medzinárodných konferenciách)

[4] Nehodiace sa prečiarknite

Declation of Honour

I hereby declare on my honour that I wrote this thesis independently under supervision of Dr. Marcel Baláž, after consultations and with use of cited literature.

Bratislava, June 2023 ....

Bc. Miroslav Hájek

Annotation

Slovak University of Technology in Bratislava

Faculty of Informatics and Information Technologies

Degree course:	Intelligent Software Systems
Author:	Bc. Miroslav Hájek
Master's Thesis:	Machinery vibrodiagnostics with the industrial internet of things
Supervisor:	Dr. Marcel Baláž
Departmental advisor:	Jakub Findura
Consultant:	Lukáš Doubravský

June 2023

Anotácia

Slovenská technická univerzita v Bratislave

Fakulta informatiky a informačných technológií

Študijný program:	Inteligentné softvérové systémy
Autor:	Bc. Miroslav Hájek
Diplomová práca:	Vibrodiagnostika strojov s priemyselným internetom vecí
Vedúci diplomovej práce:	Ing. Marcel Baláž, PhD.
Pedagogický vedúci:	Ing. Jakub Findura
Konzultant:	Ing. Lukáš Doubravský

Jún 2023

Contents

List of Figures

1 Introduction

Manufacturing is experiencing a shift in the traditional practices of asset operational status evaluation and utilization. The rise of Industry 4.0 means greater automation and robotization of the production halls to achieve optimal usage of available resources. The secondary aspect in the enterprises' endeavor, however not less important, is to keep track of the equipment wear and tear. The corrective action be it repair or replacement should be taken on time in response to the key indicators.

The goal is to preserve required safety and production efficiency while extending the useful life of mechanical rotating parts. In the factories and logistics where this sort of equipment is vital, there is a rising interest in the ability to watch in real time the machine's health status. Proactive fault diagnosis is imperative to initiate a repair without adding unnecessary costs.

Vibrations are the most non-intrusive way to sense and record eventually fatal deficiencies right at the onset. The experts use it to distinguish faulty states and to identify the malfunction's root causes. In critical circumstances such as is the case of the large turbines in the power plants, the precautions leading to regular machinery check-ups are already in place. The monitoring solution has to be sufficiently independent, reliable, and as self-sufficient as per model performance to reach wider acceptance and spread.

The main issue to consider in large-scale machinery monitoring with vibrations is that there are lots of uninformative streams of samples not directly useful for the production line operator. The dashboard must aggregate these flows into trend variables with severity levels categorized based on industrial standards. The majority of signals are viewed once at the maximum therefore to store or even transmit them from the edge device in its entirety would be wasteful. The complex overview of the mechanical equipment status is attainable only when agent devices and sensors are cheap enough with a long lifespan on battery power. Preferably they should also remain physically small to reduce the additional clutter.

Attempted machine and deep learning approaches have the crucial impediment that the construction of every single machine is unique to some extent because of tolerances and variable load. The model must be trained specifically for the target environment to achieve the ideal performance. In addition, the failures are relatively rare events that usually occur several months apart. In these circumstances, it is hard to obtain a large enough sample of fault events fast. Novelty detection is a technique that can be applied in this case.

This progress report to master's thesis is organized in the following manner. In the first chapter of problem analysis we explore the mechanical maintenance approaches and industry standards on common fault identification. Then section 2 is all about measuring vibrations and transforming them into features meaningful in automatic fault pattern recognition. In section 3 we delve into modes of diagnosis based on reduced relevant indicators. Section 4 deals with evaluation datasets used to determine computational requirements on IIoT infrastructure. Chapter 2 defines data format and proposes processing steps to diagnose the imminent failure and different fault types.

2 Problem analysis

In the problem analysis chapter we expore the feature extraction methods and machine learning algorithms for the fault diagnostics. The basis we build upon is the domain knowledge of the mechanical engineers in vibration signal measurement and its evaluation.

2.1 Condition monitoring

All rotating machinery eventually fails because of the long-term strain on the individual parts, incorrect workmanship, installation, or operational procedures. In the end, these factors cause the equipment not to fulfill its intended functionality. Many instrumentation methods are practiced to reveal evolving faults: vibration and acoustic noise monitoring, electric supply line measurements, thermography, oil and particle analysis, ultrasonic testing, etc. However, vibration signals are the preferred tool for rotating machinery monitoring [1].

The defect needs to be either repaired or replaced, preferably without significant production downtime, further damage to the other attached elements, or any endangerment of the responsible personnel. The maintenance strategies are chosen according to the machine's importance as a result of its failure effect evaluation on the system. The guide to set appropriate maintenance procedures is outlined in the IEC 60706-2 standard and involves reliability-centered maintenance analysis [2].

2.1.1 Maintenance strategies

There are three different approaches to maintenance across the industry: **reactive, preventive, and predictive** [3]. In general, the more sophisticated methods are beneficial in a high-stakes environment. The unexpected machine shutdown can have a negative economic impact on the enterprise, resulting in decreased product quality and demands spare parts to be ready in the supply inventory at all times. In certain situations it suffice to utilize a simpler maintenance program, but the predictive maintenance gains attraction in the Industry 4.0 to optimize assets' usage [4].

Reactive maintenance allows machinery to run until a complete failure. This is the most inappropriate way to maintain the production line, but it is straightforward. It requires a large stock of replacement parts on-site and breakage inflicts a 'crisis management mode' upon the plant [3]. On-demand repairs are justified if short downtime is acceptable full and swift replacement of a broken machine with a backup is possible or there is no threat of failure to the surrounding environment [5].

Preventive maintenance is performed before any issue is detected. Maintenance occurs at regular intervals derived from a predetermined period in the calendar or expected machine running time (e.g. MTTF - Mean Time To Failure). The schedule is crucial but can result in components being replaced in good condition creating waste or occasionally too late after the machine breaks. In this case conservative planning is usually the norm to keep the machine always in a perfect state and therefore more frequent intervention. [1].

Predictive maintenance known as condition-based maintenance (CbM), improves the predictability of reactive maintenance and eliminates the waste in overall resource utilization of cautious prevention. The machine downtime is scheduled after the detection of unhealthy trends in fault monitoring with sensors and troublesome components are identified.

A measurable decrease in effectivity allows us to order necessary parts in advance and organize repairs of several machines at a convenient time. The misdetection leads to increased costs compared to previous methods and raises the expectation that faults are distinguishable among themselves [6].

The P-F curve is a widespread representation of equipment degradation over time based on historical records (Fig. 2.1). Corrective action should be taken between the event of potential failure (P), when the fault detection is activated, and functional

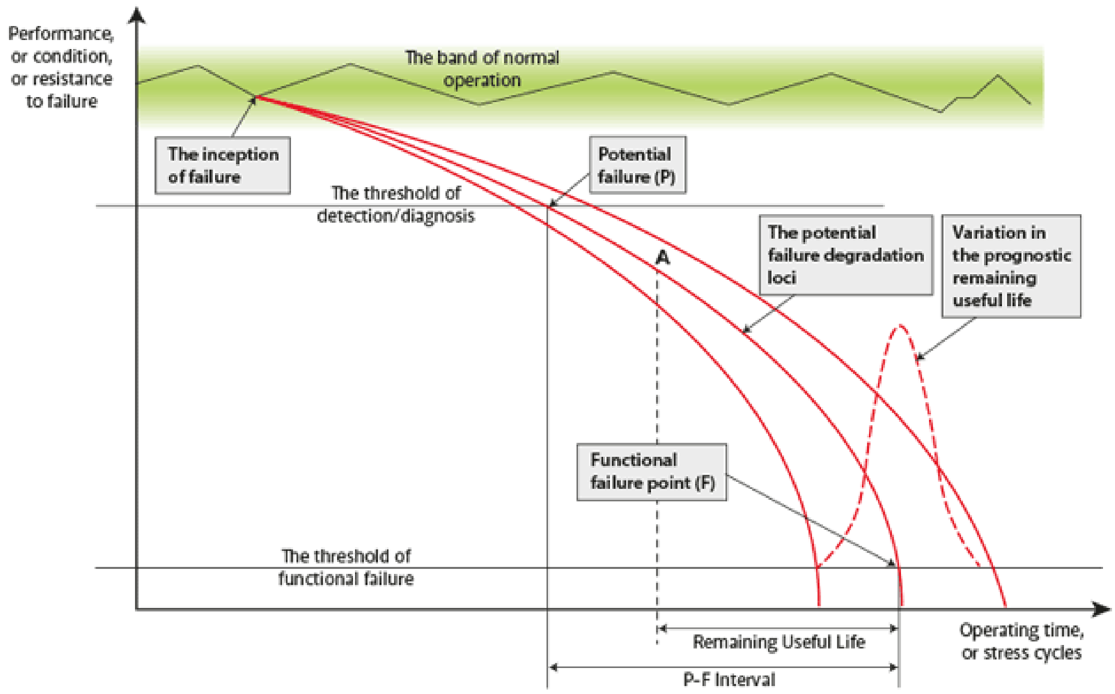


Figure 2.1: P-F curve represents the evolution of the asset's health [7]

failure point (F) in the P-F interval [8]. These division points are not exactly set but have statistical distribution to them.

The Remaining Useful Life (RUL) of the specific running machine in the given instance can be merely estimated analytically, with the survival probabilities of the individual components, and based on the model of the 'run-to-failure' histories and usage parameters [9]. Predictive condition monitoring aims to extend the machine lifespan to the maximum by predicting expected RUL.

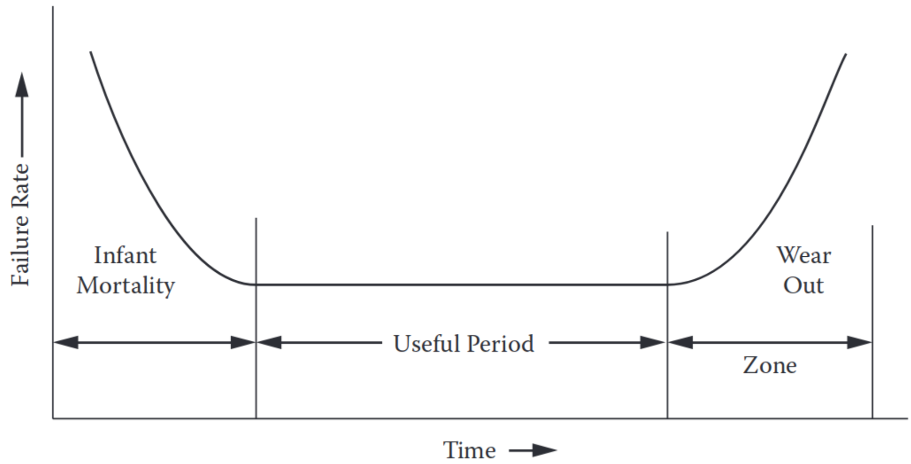


Figure 2.2: Bath tub curve [1]

A high failure rate is present not only at the worn-out stage when the parts are fatigued or corroded but also in the early stages soon after assembly. Causes can be found in manufacturing or material defects, inadequate installation, or improper start-up procedures. During the stable middle phase, malfunction can occur after the machine's excessive overload. The time plot to failure rate is known as the bath tub curve (Fig. 2.2).

2.1.2 Vibration fault types

Mechanical problems during machinery operation bring about vibrations in the vast majority of cases. Therefore vibroacoustic diagnostics is considered one of the most important methods in early component fault identification [5].

The cause of vibration comes out of the changing force in its magnitude or direction. The most emerging defects can be encompassed by explaining the deficiencies of the mechanical structure broadly categorized as **unbalance, misalignment, looseness, excentricity, deformation, crack and influence of the external force** (e.g. friction) [6]. It is important to stress that our concern are not the underlying deficiencies in mechanical parts, but the correct fault classification based on the signal waveform.

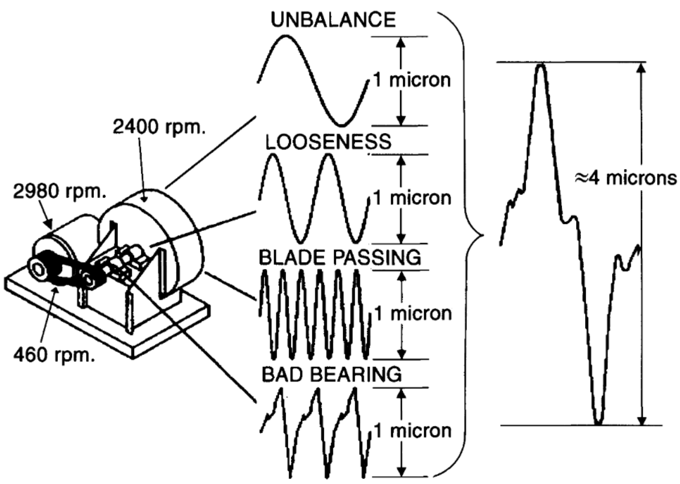


Figure 2.3: Complex machinery vibrations [6]

Rotating machine disorders do exhibit frequency signatures at various ranges in the frequency spectrum with supplementary symptoms carried in phase. Most of the occurring faults can be tied to the main rotational speed of the component under investigation (Fig. 2.3) [6]. Imbalance, misalignment, and looseness normally appear at frequencies up to 300 Hz. These low-frequency faults are associated with the movement of the shaft and primarily coincide with revolution speed and its harmonics. Bearing and gearbox defects in the late stages of development, show up in the range between 300 Hz to 1 kHz. Higher frequencies measured traditionally to a limit of 10 kHz help notice the deficiencies in bearings even sooner [10].

One of the methods vibration experts utilize in the identification of the damaged part from the frequency spectrum is **order analysis**. In essence, it is about noticing the excessive peaks at harmonic frequencies that are integer multiples of fundamental frequency (1x rpm) (Tab. 2.1):

Frequency content		**Likely reason**	**Other causes**
Synchronous	1 x rpm	Imbalance	Eccentric journals Bent shaft / Misalignment (high axial vibration) Bad belt (if rpm of belt)
	2 x rpm	Looseness	Misalignment (high axial vibration) Cracked rotor Bad belt (if rpm of belt)
	3 x rpm	Misalignment	and axial looseness
	Many x rpm	Bad gears Severe looseness	Gear teeth x rpm Fan blade count x rpm
Sub-synchronous	<1 x rpm	Oil whirl	Bad drive belt Background Resonance
Non-synchronous	-	Electrical problems (x 50 Hz) Reciprocating forces Aerodynamic forces Bad antifriction bearings	Rubbing

Table 2.1: Expert observed likely vibration causes (based on [6, 5, 11])

Because of inherent tolerances in machine manufacturing and assembly, the rotational frequency always manifests itself, even in baseline signature [6, 11]. In the most likely scenario, some faults appear as compared to rotational frequency solely in**synchronous, subsynchronous, or non-synchronous components**. The defects can occur also in a predictable combination of the ones mentioned. Other common patterns experts look for are modulation sidebands typical for bearings and gears extractable with cepstrum analysis [5]. Therefore procedure relying on elimination narrows down unrelated causes effectively.

2.1.3 Technical standards

Vibration-based condition monitoring practices adopted in the factory's predictive maintenance management must comply with normative guidelines formalized in ISO international standards. The standards are concerned with each step in the process that originates with transducer placements and data acquisition. They prescribe conventions for setting fault severity levels and provide empirically observed vibration characteristics of common defects. Two relevant standards for IoT diagnostics systems are *ISO 20816* (updated from ISO 10816) and *ISO 13373*.

ISO 20816-1:2016 establishes the approach to vibration measurement and evaluation on non-rotating housing of machinery parts [12]. The measurement units are agreed upon for kinematic quantities of vibrations. Acceleration is to be measured in meters per second squared (m/s^2), velocity in millimeters per second (mm/s), and displacement in micrometers (μm). It is customary to evaluate broad-band vibration velocity in terms of root mean square value (RMS), as it has a relation to its signal energy. No simple direct relationship is expressible among these quantities, except in stationary signals.

The vibration severity is the maximum magnitude value measured in two radial directions (horizontal, and vertical) or supplemented with a third direction along the shaft in the axial axis. Multiple measurement locations, i.e. on several bearings or couplings, should be assessed independently.

Criteria introduced to judge vibration severity are its absolute vibration magnitude, change in the magnitude vector, and rate of change. In terms of maximal magnitudes the machines of varied sizes are split into four severity zones defined in the chart (Tab. 2.2). The table values serve as guidelines towards realistic requirements between machine manufacturers and customers.

Zone A is reserved for newly commissioned machines. *Zone B* signifies suitability for long-term operation. In *zone C* is the machine deemed in unsatisfactory condition and corrective action should be taken soon. Finally, in zone D vibrations

<table>
<tr><th>Vibration velocity
RMS [mm/s]</th><th>Class I
Small machines</th><th>Class II
Medium machines</th><th>Class III
Large machines
Rigid supports</th><th>Class IV
Large machines
Flexible support</th></tr>
<tr><td>0.28</td><td rowspan="3">Good (A)</td><td rowspan="4">Good (A)</td><td rowspan="5">Good (A)</td><td rowspan="6">Good (A)</td></tr>
<tr><td>0.45</td></tr>
<tr><td>0.71</td></tr>
<tr><td>1.12</td><td rowspan="2">Satisfactory (B)</td></tr>
<tr><td>1.8</td><td rowspan="2">Satisfactory (B)</td></tr>
<tr><td>2.8</td><td rowspan="2">Unsatisfactory (C)</td><td rowspan="2">Satisfactory (B)</td></tr>
<tr><td>4.5</td><td rowspan="2">Unsatisfactory (C)</td><td rowspan="2">Satisfactory (B)</td></tr>
<tr><td>7.1</td><td rowspan="5">Unacceptable (D)</td><td rowspan="2">Unsatisfactory (C)</td></tr>
<tr><td>11.2</td><td rowspan="4">Unacceptable (D)</td><td rowspan="2">Unsatisfactory (C)</td></tr>
<tr><td>18</td><td rowspan="3">Unacceptable (D)</td></tr>
<tr><td>28</td><td rowspan="2">Unacceptable (D)</td></tr>
<tr><td>45</td></tr>
</table>

Table 2.2: ISO 20816 vibration severity chart with typical magnitudes [12]

can cause damage to the machine. The span of acceptable values differs on machine class from I through to IV with an output power of 15 kW (class I), 75 kW (class II), 10 MW (class III), or greater.

The operational limits in the form of *alarms* and *trips* are usually established on the zone boundaries or close to them. Alarms are placed between zones B and C and provide a warning about reaching the threshold significant for noticeable change. Trips in between zones C and D urge immediate action or machine shut down. Both limits should not exceed 1.25 times the upper boundary or lower zones and initially are set based on previous experience with the machine [11].

ISO 13373-1:2002 delves into further nuances of vibration monitoring and expands on procedures outlined in the vocabulary of ISO 20186. According to the standard, the data collection operates in continuous or periodic observation modes which follow an event or intervals. Both designs can be permanently mounted, but in continuous, collection 'multiplexing rate is sufficiently rapid so there is no significant data or trends lost' [11]. When channels are scanned successively with gaps between data points the system is known as 'scanning'.

The condition monitoring programme is run according to a flowchart adapted from one designed in the standard specifically to best benefit the plant. Those steps can be summarized as follows [11]:

1. Review machinery history and establish failure modes.
2. When vibration monitoring is not applicable check for other condition moni-

toring techniques or resort to preventive maintenance.

3. Select monitoring points and take preliminary vibration measurements.
4. Select vibration monitoring techniques: broadband, frequency analysis, or special techniques, and set parameters of measurement units.
5. Take baseline measurements.
6. Change levels that would warrant investigation.
7. Carry out routine condition monitoring.
8. If an alarm was exceeded, notify appropriate personnel to review data and trends, perform diagnostic evaluation, and repair as necessary. In case a new baseline is needed continue in the step of taking baseline measurements.
9. Shut down the machine when the trip level is exceeded. Then proceed the same as after the alarm trigger.

Measurement of vibrations should be accompanied by a description of the machine and its operating conditions. The machine description includes the machine identifier and its type, power source, rated rotation speed and power, configuration (shaft or belt driven), and machine support. Measurement parameters such as timestamp, transducer type, sensor location and orientation in MIMOSA code, measurement units and units qualifier (p-p, RMS), and other processing options (filters, number of averages, etc.) are to be recorded alongside the measurement value itself [11].

The transducer of choice for condition monitoring is the accelerometer which can provide the acceleration value of the body and velocity after signal integration. However, standard cautions against double integrating for displacement. The recommended frequency range for an accelerometer is 0.1 Hz to 30 kHz. The choice of transducer mount significantly lowers its resonance frequency which is least influenced by stud mount and stiff cement mount. The resonance is reduced to around 8 kHz with the use of soft epoxy or permanent magnet.

Broadband measurement requires ‘frequency ranges of 0.2 times the lowest rotational frequency to the highest frequency of interest’ [11], not exceeding 10 kHz, with

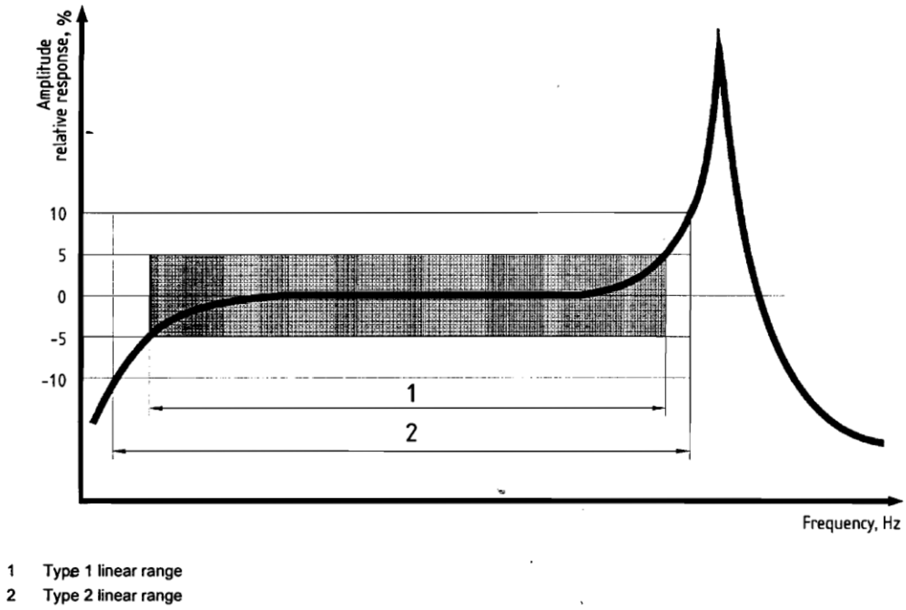


Figure 2.4: The transducer linear response and resonance in tolerance intervals [11]

RMS velocity 0.1 - 100 mm/s. Bearings and gears diagnosis may push the upper-frequency limit even higher. The tolerances of amplitude and frequency calibrations fall into two types with allowable tolerances of $\pm 5\%$ or $\pm 10\%$ (Fig. 2.4).

Equipment's 'health' can be mischaracterized when there are significant differences in the machine's normal operating conditions. Baseline measurements in all acceptable conditions are to be acquired to reduce the error in vibration evaluation. According to bath tub curve (Fig. 2.2) reference signatures should be obtained after the initial part wear-in period. The reference spectral mask of the baseline condition is designed if maximal acceptance amplitudes are different for each significant frequency band [5].

The vibration baseline is defined by broad-band magnitudes and phases of motion vectors, waveforms in time and frequency domain, the rotational speed of the machine as well as its frequency response to different speeds during start-up and coast-down captured in the Bode plot and waterfall plot. Changes during the machine operation are then depicted in value trends. Trends can be shown of overall amplitudes or limited to frequency bands.

2.2 Feature engineering

Large domain knowledge with compared to other areas of machine learning (mechanics - physics)

2.2.1 Preprocessing

- Detrending - DC removal filter
- Adaptive noise canellation of the background noise
- Time synchronous averaging

2.2.2 Feature extraction

[13] [14] [15] [16] [17] [18] [19] [20] [21] [22]
[23]

Statistical measures

- Standard Deviation
- Max. amplitude
- RMS amplitude
- Skewness
- Kurtosis

  —

- Spectral centroid
- RMS frequency
- Root variance frequency
- Spectral kurtosis / Fast kurtogram
- Harmonics (peaks) [24] [25] [26] [27] [28] [29]

- Spectral Envelope
- Harmonic spectral deviation

  —

- Energy
- Spectral negentropy [30]
- TKEO - Teager-Kaiser energy operator [31]

Signal decompositions - sparse approximations Matching pursuit algorithm optimalization problem [32]

- FFT - Short Time Fourier Transform with Hamming window and Welch averaging
- CWT-SST - Synchrosqueezing Wavelet Transform (vs. Transient-extracting transform) [33] [34] [35]
- WPD - Wavelet Packet Decomposition - to approximation and detail coef. (Fejer-Korovkin wavelet) [36] [37]
- EWT - Empirical Wavelet Transform - (Meyer wavelet) [38] [39] [32] [40] [41] [42] [43] [44]

2.2.3 Feature transformation

[13]

- Principal Component Analysis (PCA)
- Log transformation (Box-Cox Transform) to normal distribution
- Normalization (min-max, standardize)

2.2.4 Feature selection

[22] Filter method - SelectKBest in evaluation phase

- Variance Threshold
- Pearson correlation
- ANOVA F-value
- Mutal information
- Fisher score
- Spectral feature selection algorithm (SPEC)

2.3 Diagnostics techniques

Idenification of faulty states in data streams in semi-supervised learning

[45]

2.3.1 Novelty detection

[46]

- Local Outlier Factor, Local Correlation Integral (Anomaly score) [47] [48]
- DenStream (Density based clustering - DBSCAN) [49] [50] [**aggarwal_data_2014**] [51] [52]
- Half-space Trees (Isolation forest) [53] [46]

2.3.2 Classification

- kNN + Metric Tree (M-Tree for neighbourhood queries) + Euclidian Mahalanobis distance / RBF similarity

  [54] [55] [56]

[57] [58] [59] [60]

[61]

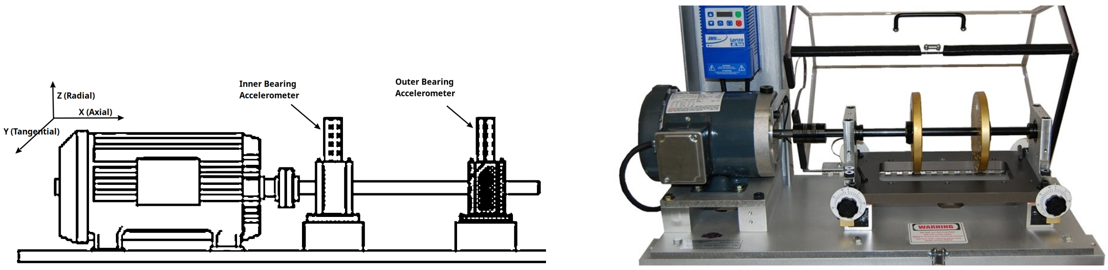
2.4 Evaluation datasets

The experimentally designed features' relevancy is first proven in comparison to comprehensive benchmark datasets. There are a few standardized datasets used in the related work, e.g. [63]. The datasets listed below are also publically available in Comma-Separated Values files online.

MaFaulDa dataset combines vibration and acoustic measurements of the shaft in deviating positions and bearings abnormalities. *CWRU dataset* focuses solely on faults in ball bearings. Another less known dataset concerns shaft unbalance, but compared to the previous two, it demonstrates behavior during revolution speed up.

2.4.1 Machinery Fault Database

MaFaulDa[1] is a collection of 1951 multivariate time series for 4 different operational conditions on rotor kit Alignment Balance Vibration Trainer (ABVT) (Fig. 2.5). Each series has 5 seconds in duration and is captured at 50 kHz. Vibration signals were obtained with piezoelectric accelerometers with a linear response up to 10 kHz, amplitude range to $\pm 490\ m/s^2$, and resolution step of 10.2 mV per m/s^2.



(a) Schematic diagram [62]

(b) Mechanical construction [64]

Figure 2.5: Machinery fault simulator for MaFaulDa

Observations were conducted in three cardinal axes simultaneously with 2 sets of accelerometers each one associated with one bearing (inner and outer bearings) (Fig. 2.5). Additionally, a magnetic tachometer produced a pulse on shaft turn. The cardioid condenser microphone recorded sound emissions with a frequency range 20 Hz - 20 kHz. Sensors were fed into a four-channel dynamic signal acquisition module.

[1]`https://www02.smt.ufrj.br/~offshore/mfs/page_01.html`

Columns in the dataset are organized as depicted in table 2.3. Machine rotational speeds were kept constant during a particular measurement, but covered a range from 737 to 3686 rpm with steps of approximately 60 rpm (equiv. 10 Hz - 60 Hz) [62]. The maximal rotational frequency achieved with a high unbalance load is 3300 rpm.

Columns	**Description**
1.	Pulse with modulation of tachometer signal to estimate rotation frequency (in TTL levels)
2., 3., 4.	Underhang bearing accelerometer (inner - between the rotor and motor) - axial, radial, tangential direction
5., 6., 7.	Overhang bearing accelerometer (Outer - outside most position after the rotor) - axial, radial, tangential direction
8.	Microphone

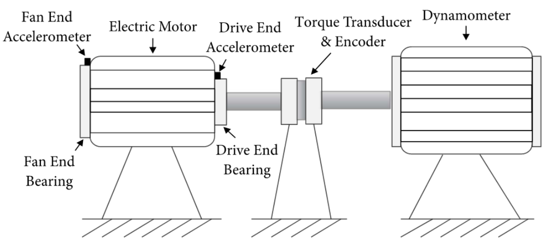
Table 2.3: MaFaulDa description of columns

This database contains normal operating conditions, faults out of unbalance, horizontal and vertical shaft misalignment, and three types of faulty bearings in inner and outer positions: outer track, inner track, rolling elements [62].

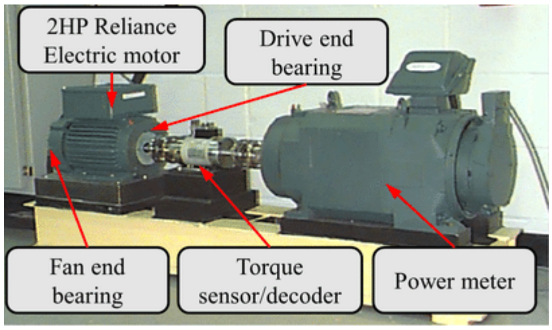
- **Normal** conditions are baseline without the adverse effect of fault in 49 different rotation speeds.
- **Unbalance** shaft time series uses 8 unbalancing weights from 6 to 35 grams and varying 45 - 49 speeds for each weight adding to 333 mass unbalance loads.
- **Vertical misalignement** set is comprised of 50 signals each (or 51 in one instance) obtained under displacements: 0.51, 0.63, 1.40, 1.90, 1.27, 1.78 mm.
- **Horizontal misalignement** signals were recorded under displacements: 0.50, 1.00, 1.50, 2.00 mm, each with 49 different speeds (or 50 in one instance) [62].
- **Bearing faults** are unnoticable without unbalance. Therefore, weights of 6, 20, and 35 grams were attached to induce a detectable effect. Each unbalance mass was combined with cage, outer race, and ball faults at multiple rotation speeds, usually at 50 different speeds.

2.4.2 CWRU bearings dataset

In Case Western Reserve University (CWRU) bearing dataset[2] recordings were made of a fan end and drive end bearings under motor loads of 0, 1, 2, and 3 Horsepower (equivalently 0, 0.75, 1.49, 2.24 kW). Shaft speed was unaltered in all experiments, but it fluctuated between 1720 and 1797 rpm (approx. 29 Hz).



(a) Schematic diagram [65]



(b) Mechanical construction [66]

Figure 2.6: CWRU machine apparatus

Single point defects were created with diameters of 0.007, 0.014, 0.021, 0.028, and 0.040 inches (equivalently 0.18, 0.36, 0.72, 1.02 mm). Fault locations on bearings are in the inner raceway, in the outer raceway directly and orthogonally relative to the load zone, and on rolling ball elements (Fig. 2.6) [57].

Columns	**Description**
1. DE	Drive end accelerometer samples
2. FE	Fan end accelerometer samples
3. BA	Base accelerometer samples (optional)
4. RPM	Rotation speed of the motor in rpm

Table 2.4: CWRU dataset description of columns

The sampling frequency during baseline set, drive end, and fan end bearing capture is 12 kHz, exclusively for drive end bearings samples were taken at 48 kHz. The duration of the time series is varied from 5 to 40 seconds. Drive end and fan end bearings signals are measured in each experiment sometimes accelerometer was mounted on the supporting base plate.

[2] `https://engineering.case.edu/bearingdatacenter/download-data-file`

2.4.3 Unbalance on the rotating shaft

Unbalance Detection of a Rotating Shaft[3] is a Kaggle dataset that simulates 4 different unbalance strengths.

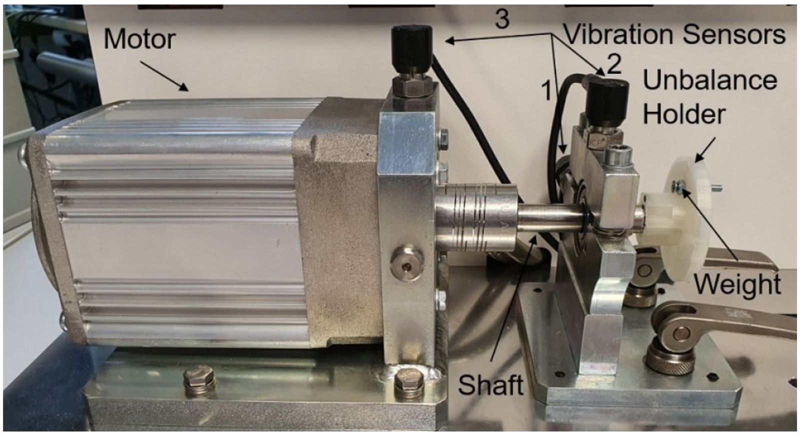


Figure 2.7: Motor driving shaft in unbalance measurement [67]

The setup is shown in Fig. 2.7. Mass of 3.28 grams (or 6.61 grams during severe unbalance test) is attached to unbalance holder successively in 5 sets (numbered 0 - 4) on the radii 0, 14, 18.5, 23, 23 mm. The rotation speed of the motor is perpetually rising between 630 and 2330 rpm in development datasets (marked with suffix D) and speeds from 1060 to 1900 rpm in the evaluation datasets (suffix E). The vibrations were recorded at a sampling rate of 4 kHz [67].

Columns	**Description**
1. V_in	Input voltage to the motor controller (V)
2. Measured_RPM	Rotation speed of the motor (rpm)
3. Vibration_1	1. Vibration sensor (samples)
4. Vibration_2	2. Vibration sensor (samples)
5. Vibration_3	3. Vibration sensor (samples)

Table 2.5: 'Unbalance on the rotating shaft' dataset description of columns

The accelerometers used are piezoelectric and have a frequency range of up to 10 KHz, dynamic range of ± 490 m/s^2, and resolution step of 10.2 mV per m/s^2. These sensor parameters are the same as in the case of MaFaulDa. In total, three different uniaxial accelerometers are mounted on the motor housing.

[3] https://www.kaggle.com/datasets/jishnukoliyadan/vibration-analysis-on-rotating-shaft

3 Design

3.1 Research questions

1. *Which time-frequency features can be extracted from vibrational signals to provide an accurate record of machinery faults?*

2. *What are the savings in transmission bandwidth when chosen signal features are used in comparison to raw sampled measurement or lossless compression techniques?*

3. *How can the machinery faults be continuously identified based on collected events?*

3.2 Infrastructure

- **Input:** Samples from three-axis MEMS accelerometers, RPM tachometer, Noise background

- **Output is either:** machine overall status, type of fault,

- **Output for domain expert**: Control chart of trend features

1. MEMS accelerometers are placed on at least two distinict measurement points in two perpendicular axis and one sensor in base for denoising. Rotational speed is captured at the same time too.

2. Sensors are triggered in regular intervals (every 15 minutes) to collect sample recording from the band saw.

3. **Features** are computed and compared to recent measurements. If there is an statistically significant change the whole summary is send, otherwise keepalive notification is send.

4. Set of features - at the explatory stage all are computed and stored locally. Use features in conjunction with transformation techniques and diagnostics model to reduce the set of features to minimum with the k-fold cross-validation.

5. Database stores history of measurements

6. **Diagnosis panel runs clustering** with introduction of annotations to notify the operator about observed fault and imminent failure of the machine.

Literature

1. MOHANTY, Amiya Ranjan. *Machinery Condition Monitoring: Principles and Practices*. CRC Press, 2015. ISBN 978-1-4665-9305-3.
2. EL-THALJI, Idriss. Predictive maintenance (PdM) analysis matrix: A tool to determine technical specifications for PdM ready-equipment. *IOP Conference Series: Materials Science and Engineering*. 2019, vol. 700, p. 012033. Available from DOI: 10.1088/1757-899X/700/1/012033.
3. SCHEFFER, C.; GIRDHAR, P. *Practical Machinery Vibration Analysis and Predictive Maintenance*. IDC Technologies, Elsevier, 2004. ISBN 0-7506-6275-1.
4. ÇINAR, Zeki Murat; ABDUSSALAM NUHU, Abubakar; ZEESHAN, Qasim; KORHAN, Orhan; ASMAEL, Mohammed; SAFAEI, Babak. Machine Learning in Predictive Maintenance towards Sustainable Smart Manufacturing in Industry 4.0. *Sustainability* [online]. 2020, vol. 12, no. 19, p. 8211 [visited on 2023-04-23]. ISSN 2071-1050. Available from DOI: 10.3390/su12198211. Number: 19 Publisher: Multidisciplinary Digital Publishing Institute.
5. ŽIARAN, Stanislav. *Technická diagnostika*. 1. vyd. Bratislava: Vydavateľstvo STU, 2013. ISBN 978-80-227-4051-7.
6. DAVIES, A. *Handbook of Condition Monitoring: Techniques and Methodology*. Dordrecht: Springer Netherlands, 2012. ISBN 978-94-011-4924-2. OCLC: 958541223.
7. JENNIONS, Ian K. *Integrated Vehicle Health Management: Perspectives on an Emerging Field*. SAE International, 2011. ISBN 978-0-7680-6432-2. Google-Books-ID: PIi7pwAACAAJ.

8. BOUSDEKIS, Alexandros; MENTZAS, Gregoris. Enterprise Integration and Interoperability for Big Data-Driven Processes in the Frame of Industry 4.0. *Frontiers in Big Data*. 2021, vol. 4. Available from DOI: 10.3389/fdata.2021.644651.

9. OKOH, C.; ROY, R.; MEHNEN, J.; REDDING, L. Overview of Remaining Useful Life Prediction Techniques in Through-life Engineering Services. *Procedia CIRP* [online]. 2014, vol. 16, pp. 158–163 [visited on 2023-04-23]. ISSN 2212-8271. Available from DOI: 10.1016/j.procir.2014.02.006.

10. TORRES, Pedro; RAMALHO, Armando; CORREIA, Luis. Automatic Anomaly Detection in Vibration Analysis Based on Machine Learning Algorithms. In: MACHADO, José; SOARES, Filomena; TROJANOWSKA, Justyna; YILDIRIM, Sahin; VOJTĚŠEK, Jiří; REA, Pierluigi; GRAMESCU, Bogdan; HRYBIUK, Olena O. (eds.). *Innovations in Mechatronics Engineering II*. Cham: Springer International Publishing, 2022, pp. 13–23. Lecture Notes in Mechanical Engineering. ISBN 978-3-031-09385-2. Available from DOI: 10.1007/978-3-031-09385-2_2.

11. *ISO 13373-1:2002 - Condition monitoring and diagnostics of machines - Vibration condition monitoring - Part 1: General procedures*. International Organization for Standardization, 2002. Available also from: https://www.iso.org/standard/21831.html.

12. *ISO 20816-1:2016 - Mechanical vibration - Measurement and evaluation of machine vibration - Part 1: General guidelines* [online]. International Organization for Standardization, 2016 [visited on 2022-11-14]. Available from: https://www.iso.org/standard/63180.html.

13. ZHENG, Alice; CASARI, Amanda. *Feature Engineering for Machine Learning*. O'Reilly Media, 2018. ISBN 978-1-4919-5324-2.

14. JOHNSON, Max Kuhn {and} Kjell. *Feature Engineering and Selection: A Practical Approach for Predictive Models* [online]. 2019. [visited on 2023-02-26]. Available from: http://www.feat.engineering/.

15. ALTAF, Muhammad; AKRAM, Tallha; KHAN, Muhammad Attique; IQBAL, Muhammad; CH, M. Munawwar Iqbal; HSU, Ching-Hsien. A New Statistical Features Based Approach for Bearing Fault Diagnosis Using Vibration Signals. *Sensors* [online]. 2022, vol. 22, no. 5, p. 2012 [visited on 2023-03-27]. ISSN 1424-8220. Available from DOI: `10.3390/s22052012`. Number: 5 Publisher: Multidisciplinary Digital Publishing Institute.

16. MOSTAFAVI, Alireza; SADIGHI, Ali. A Novel Online Machine Learning Approach for Real-Time Condition Monitoring of Rotating Machines. In: *2021 9th RSI International Conference on Robotics and Mechatronics (ICRoM)*. 2021, pp. 267–273. Available from DOI: `10.1109/ICRoM54204.2021.9663495`. ISSN: 2572-6889.

17. BRITO, Lucas Costa; SUSTO, Gian Antonio; BRITO, Jorge Nei; DUARTE, Marcus Antonio Viana. Fault Detection of Bearing: An Unsupervised Machine Learning Approach Exploiting Feature Extraction and Dimensionality Reduction. *Informatics* [online]. 2021, vol. 8, no. 4, p. 85 [visited on 2023-03-27]. ISSN 2227-9709. Available from DOI: `10.3390/informatics8040085`. Number: 4 Publisher: Multidisciplinary Digital Publishing Institute.

18. PEETERS, Geoffroy. A large set of audio features for sound description. 2004.

19. ZHUO, Rongjin; DENG, Zhaohui; CHEN, Bing; LIU, Tao; GE, Jimin; LIU, Guoyue; BI, Shenghao. Research on online intelligent monitoring system of band saw blade wear status based on multi-feature fusion of acoustic emission signals. *The International Journal of Advanced Manufacturing Technology* [online]. 2022, vol. 121, no. 7, pp. 4533–4548 [visited on 2023-02-11]. ISSN 1433-3015. Available from DOI: `10.1007/s00170-022-09515-3`.

20. MOHAMMADI, Hamed. *Early Detection of Imbalance in Load and Machine In Front Load Washing Machines by Monitoring Drum Movement*. 2020. Master's Thesis.

21. EGAJI, Oche A; EKWEVUGBE, Tobore; GRIFFITHS, Mark. A Data Mining based Approach for Electric Motor Anomaly Detection Applied on Vibration Data. In: *2020 Fourth World Conference on Smart Trends in Systems, Security*

and Sustainability (WorldS4). 2020, pp. 330–334. Available from DOI: 10.1109/WorldS450073.2020.9210318.

22. NANDI, Asoke Kumar; AHMED, Hosameldin. *Condition monitoring with vibration signals: compressive sampling and learning algorithms for rotating machines*. Hoboken, NJ, USA: Wiley-IEEE Press, 2019. ISBN 978-1-119-54462-3.

23. JUNG, Deokwoo; ZHANG, Zhenjie; WINSLETT, Marianne. Vibration Analysis for IoT Enabled Predictive Maintenance. In: *2017 IEEE 33rd International Conference on Data Engineering (ICDE)*. 2017, pp. 1271–1282. Available from DOI: 10.1109/ICDE.2017.170. ISSN: 2375-026X.

24. HATEM, G. M.; SADAH, J. W. Abdul; SAEED, T. R. Comparative Study of Various CFAR Algorithms for Non-Homogenous Environments. *IOP Conference Series: Materials Science and Engineering* [online]. 2018, vol. 433, no. 1, p. 012080 [visited on 2022-12-09]. ISSN 1757-899X. Available from DOI: 10.1088/1757-899X/433/1/012080. Publisher: IOP Publishing.

25. BISHOP, Steven M.; ERCOLE, Ari. Multi-Scale Peak and Trough Detection Optimised for Periodic and Quasi-Periodic Neuroscience Data. In: HELDT, Thomas (ed.). *Intracranial Pressure & Neuromonitoring XVI*. Cham: Springer International Publishing, 2018, pp. 189–195. Acta Neurochirurgica Supplement. ISBN 978-3-319-65798-1. Available from DOI: 10.1007/978-3-319-65798-1_39.

26. ADIKARAM, K.K. Lasantha Britto; HUSSEIN, Mohamed; EFFENBERGER, Mathias; BECKER, T. Non-Parametric Local Maxima and Minima Finder with Filtering Techniques for Bioprocess. *Journal of Signal and Information Processing*. 2016, vol. 07, pp. 192–213. Available from DOI: 10.4236/jsip.2016.74018.

27. GERBER, Timothée; MARTIN, Nadine; MAILHES, Corinne. Identification of harmonics and sidebands in a finite set of spectral components. 2013, vol. 1.

28. NUNES, Leonardo; ESQUEF, Paulo Antonio; BISCAINHO, Luiz. Evaluation of Threshold-Based Algorithms for Detection of Spectral Peaks in Audio. In: 2007.

29. KARLÖF, L.; ØLGÅRD, T. A.; GODTLIEBSEN, F.; KACZMARSKA, M.; FISCHER, H. Statistical techniques to select detection thresholds for peak signals in ice-core data. *Journal of Glaciology* [online]. 2005, vol. 51, no. 175, pp. 655–662 [visited on 2022-08-27]. ISSN 0022-1430, ISSN 1727-5652. Available from DOI: `10.3189/172756505781829115`. Publisher: Cambridge University Press.

30. AVOCI, Moise. Spectral negentropy and kurtogram performance comparison for bearing fault diagnosis. In: Dubrovnik, Croatia: International Measurement Confederation (IMEKO), 2020.

31. SHI, Xiangfu; ZHANG, Zhen; XIA, Zhiling; LI, Binhua; GU, Xin; SHI, Tingna. Application of Teager–Kaiser Energy Operator in the Early Fault Diagnosis of Rolling Bearings. *Sensors* [online]. 2022, vol. 22, no. 17, p. 6673 [visited on 2023-04-02]. ISSN 1424-8220. Available from DOI: `10.3390/s22176673`.

32. SONG, Yongxing; LIU, Jingting; WU, Dazhuan; ZHANG, Linhua. The MFBD: a novel weak features extraction method for rotating machinery. *Journal of the Brazilian Society of Mechanical Sciences and Engineering* [online]. 2021, vol. 43, no. 12, p. 547 [visited on 2023-03-27]. ISSN 1806-3691. Available from DOI: `10.1007/s40430-021-03259-z`.

33. ARTS, Lukas P. A.; BROEK, Egon L. van den. The fast continuous wavelet transformation (fCWT) for real-time, high-quality, noise-resistant time–frequency analysis. *Nature Computational Science* [online]. 2022, vol. 2, no. 1, pp. 47–58 [visited on 2022-08-27]. ISSN 2662-8457. Available from DOI: `10.1038/s43588-021-00183-z`. Number: 1 Publisher: Nature Publishing Group.

34. YU, Gang. A Concentrated Time–Frequency Analysis Tool for Bearing Fault Diagnosis. *IEEE Transactions on Instrumentation and Measurement*. 2020, vol. 69, no. 2, pp. 371–381. ISSN 1557-9662. Available from DOI: `10.1109/TIM.2019.2901514`. Conference Name: IEEE Transactions on Instrumentation and Measurement.

35. HERRERA, Roberto; BAAN, Mirko; HAN, Jiajun. Applications of the synchrosqueezing transform in seismic time-frequency analysis. *Geophysics*. 2014, vol. 79, pp. V55–V64. Available from DOI: `10.1190/geo2013-0204.1`.

36. YEN, Gary; LIN, K.C. Wavelet Packet Feature Extraction for Vibration Monitoring. *Industrial Electronics, IEEE Transactions on*. 2000, vol. 47, pp. 650–667. Available from DOI: 10.1109/41.847906.

37. WICKMANN, Jon. *A wavelet approach to dimension reduction and classification of hyperspectral data*. Oslo, 2007. Master's Thesis. Faculty of Mathematics and Natural Sciences, University of Oslo.

38. WANG, Yung-Hung; YEH, Chien-Hung; YOUNG, Hsu-Wen Vincent; HU, Kun; LO, Men-Tzung. On the computational complexity of the empirical mode decomposition algorithm. *Physica A: Statistical Mechanics and its Applications* [online]. 2014, vol. 400, pp. 159–167 [visited on 2022-09-01]. ISSN 0378-4371. Available from DOI: 10.1016/j.physa.2014.01.020.

39. TIWARI, Prashant; UPADHYAY, S. H. Novel self-adaptive vibration signal analysis: Concealed component decomposition and its application in bearing fault diagnosis. *Journal of Sound and Vibration* [online]. 2021, vol. 502, p. 116079 [visited on 2022-08-27]. ISSN 0022-460X. Available from DOI: 10.1016/j.jsv.2021.116079.

40. LI, Zheng; MING, Anbo; ZHANG, Wei; LIU, Tao; CHU, Fulei; LI, Yin. Fault Feature Extraction and Enhancement of Rolling Element Bearings Based on Maximum Correlated Kurtosis Deconvolution and Improved Empirical Wavelet Transform. *Applied Sciences*. 2019, vol. 9, p. 1876. Available from DOI: 10.3390/app9091876.

41. ZHUANG, Cuifang; LIAO, Ping. An Improved Empirical Wavelet Transform for Noisy and Non-Stationary Signal Processing. *IEEE Access*. 2020, vol. PP, pp. 1–1. Available from DOI: 10.1109/ACCESS.2020.2968851.

42. YONGGANG, xu; LI, Shuang; TIAN, Weikang; YU, Jun; ZHANG, Kun. Time and frequency domain scanning fault diagnosis method based on spectral negentropy and its application. *The International Journal of Advanced Manufacturing Technology*. 2020, vol. 108. Available from DOI: 10.1007/s00170-020-05302-0.

43. XU, Yonggang; ZHANG, Kun; MA, Chaoyong; SHENG, Zhipeng; SHEN, Hongchen. An Adaptive Spectrum Segmentation Method to Optimize Empirical Wavelet Transform for Rolling Bearings Fault Diagnosis. *IEEE Access*. 2019, vol. 7, pp. 30437–30456. ISSN 2169-3536. Available from DOI: `10.1109/ACCESS.2019.2902645`. Conference Name: IEEE Access.

44. NI, Ruizheng; QIU, Ruichang; JIN, Zheming; CHEN, Jie; LIU, Zhigang. Improved empirical wavelet transform (EWT) and its application in non-stationary vibration signal of transformer. *Scientific Reports* [online]. 2022, vol. 12, no. 1, p. 17533 [visited on 2023-04-24]. ISSN 2045-2322. Available from DOI: `10.1038/s41598-022-22519-z`. Number: 1 Publisher: Nature Publishing Group.

45. YU, Wennian; KIM, Il Yong; MECHEFSKE, Chris. Analysis of different RNN autoencoder variants for time series classification and machine prognostics. *Mechanical Systems and Signal Processing* [online]. 2021, vol. 149, p. 107322 [visited on 2022-11-12]. ISSN 0888-3270. Available from DOI: `10.1016/j.ymssp.2020.107322`.

46. TOGBE, Maurras Ulbricht; BARRY, Mariam; BOLY, Aliou; CHABCHOUB, Yousra; CHIKY, Raja; MONTIEL, Jacob; TRAN, Vinh-Thuy. Anomaly Detection for Data Streams Based on Isolation Forest Using Scikit-Multiflow. In: GERVASI, Osvaldo; MURGANTE, Beniamino; MISRA, Sanjay; GARAU, Chiara; BLEČIĆ, Ivan; TANIAR, David; APDUHAN, Bernady O.; ROCHA, Ana Maria A. C.; TARANTINO, Eufemia; TORRE, Carmelo Maria; KARACA, Yeliz (eds.). *Computational Science and Its Applications – ICCSA 2020* [online]. Cham: Springer International Publishing, 2020, vol. 12252, pp. 15–30 [visited on 2023-03-24]. ISBN 978-3-030-58810-6 978-3-030-58811-3. Available from DOI: `10.1007/978-3-030-58811-3_2`. Series Title: Lecture Notes in Computer Science.

47. JANSSENS, Jeroen; POSTMA, Eric. One-class classification with LOF and LOCI: An empirical comparison. 2007.

48. YU, Kangqing; SHI, Wei; SANTORO, Nicola. Designing a Streaming Algorithm for Outlier Detection in Data Mining - An Incremental Approach. *Sensors* [online]. 2020, vol. 20, no. 5, p. 1261 [visited on 2023-03-24]. ISSN 1424-8220. Available from DOI: 10.3390/s20051261. Number: 5 Publisher: Multidisciplinary Digital Publishing Institute.

49. CAO, Feng; ESTERT, Martin; QIAN, Weining; ZHOU, Aoying. Density-Based Clustering over an Evolving Data Stream with Noise. In: *Proceedings of the 2006 SIAM International Conference on Data Mining* [online]. Society for Industrial and Applied Mathematics, 2006, pp. 328–339 [visited on 2023-03-24]. ISBN 978-0-89871-611-5 978-1-61197-276-4. Available from DOI: 10.1137/1.9781611972764.29.

50. SHUKLA, M.; KOSTA, Y. P.; JAYSWAL, M. A Modified Approach of OPTICS Algorithm for Data Streams. *Engineering, Technology & Applied Science Research* [online]. 2017, vol. 7, no. 2, pp. 1478–1481 [visited on 2023-03-24]. ISSN 1792-8036. Available from DOI: 10.48084/etasr.963.

51. GHESMOUNE, Mohammed; LEBBAH, Mustapha; AZZAG, Hanene. State-of-the-art on clustering data streams. *Big Data Analytics* [online]. 2016, vol. 1, no. 1, p. 13 [visited on 2023-03-24]. ISSN 2058-6345. Available from DOI: 10.1186/s41044-016-0011-3.

52. ZHAO, Yikai; ZHONG, Zheng; LI, Yuanpeng; ZHOU, Yi; ZHU, Yifan; CHEN, Li; WANG, Yi; YANG, Tong. Cluster-Reduce: Compressing Sketches for Distributed Data Streams. In: *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining* [online]. Virtual Event Singapore: ACM, 2021, pp. 2316–2326 [visited on 2023-02-11]. ISBN 978-1-4503-8332-5. Available from DOI: 10.1145/3447548.3467217.

53. TAN, Swee Chuan; TING, Kai Ming; LIU, Tony Fei. Fast Anomaly Detection for Streaming Data. 2011.

54. SHENG, Hao; CHEN, Zhongsheng; XIA, Yemei; HE, Jing. Review of Artificial Intelligence-based Bearing Vibration Monitoring. In: *2020 11th International Conference on Prognostics and System Health Management (PHM-2020 Ji-*

nan). 2020, pp. 58–67. Available from DOI: 10.1109/PHM-Jinan48558.2020.00018. ISSN: 2166-5656.

55. WAGNER, Tal; GUHA, Sudipto; KASIVISWANATHAN, Shiva Prasad; MISHRA, Nina. Semi-Supervised Learning on Data Streams via Temporal Label Propagation. In: *Proceedings of the 35th International Conference on Machine Learning*. 2018.

56. HUBERT, Mia; DEBRUYNE, Michiel; ROUSSEEUW, Peter J. Minimum covariance determinant and extensions. *WIREs Computational Statistics* [online]. 2018, vol. 10, no. 3, e1421 [visited on 2023-03-24]. ISSN 1939-0068. Available from DOI: 10.1002/wics.1421. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002

57. JAMIL, Mohd Atif; KHAN, Md Asif Ali; KHANAM, Sidra. Feature-based performance of SVM and KNN classifiers for diagnosis of rolling element bearing faults. *Vibroengineering PROCEDIA* [online]. 2021, vol. 39, pp. 36–42 [visited on 2023-04-07]. ISSN 2345-0533. Available from DOI: 10.21595/vp.2021.22307. Publisher: JVE International Ltd.

58. GOUMAS, Stefanos; ZERVAKIS, Michalis; STAVRAKAKIS, G. Classification of washing machines vibration signals using discrete wavelet analysis for feature extraction. *Instrumentation and Measurement, IEEE Transactions on*. 2002, vol. 51, pp. 497–508. Available from DOI: 10.1109/TIM.2002.1017721.

59. CHAPELLE, Olivier; SCHÖLKOPF, Bernhard; ZIEN, Alexander (eds.). *Semi-supervised learning*. Cambridge, Mass: MIT Press, 2006. Adaptive computation and machine learning. ISBN 978-0-262-03358-9. OCLC: ocm64898359.

60. DOBILAS, Saul. *Semi-Supervised Learning — How to Assign Labels with Label Propagation Algorithm* [Medium] [online]. 2022-02-05. [visited on 2023-03-24]. Available from: https://towardsdatascience.com/semi-supervised-learning-how-to-assign-labels-with-label-propagation-algorithm-9f1683f4d0eb.

61. MAURYA, Seetaram; SINGH, Vikas; VERMA, Nishchal K.; MECHEFSKE, Chris K. Condition-Based Monitoring in Variable Machine Running Conditions Using Low-Level Knowledge Transfer With DNN. *IEEE Transactions on*

Automation Science and Engineering. 2021, vol. 18, no. 4, pp. 1983–1997. ISSN 1558-3783. Available from DOI: 10.1109/TASE.2020.3028151. Conference Name: IEEE Transactions on Automation Science and Engineering.

62. PESTANA-VIANA, Denys; ZAMBRANO-LOPEZ, Rafael; LIMA, Amaro A. de; DE M. PREGO, Thiago; NETTO, Sergio L.; SILVA, Eduardo A.B. da. The influence of feature vector on the classification of mechanical faults using neural networks. In: *2016 IEEE 7th Latin American Symposium on Circuits & Systems (LASCAS)* [online]. Florianopolis: IEEE, 2016, pp. 115–118 [visited on 2023-04-28]. ISBN 978-1-4673-7835-2. Available from DOI: 10.1109/LASCAS.2016.7451023.

63. RIBEIRO, Felipe; MARINS, Matheus; NETTO, Sergio; SILVA, Eduardo. Rotating machinery fault diagnosis using similarity-based models. In: *Anais de XXXV Simpósio Brasileiro de Telecomunicações e Processamento de Sinais* [online]. Sociedade Brasileira de Telecomunicações, 2017 [visited on 2023-04-28]. Available from DOI: 10.14209/sbrt.2017.133.

64. *SpectraQuest Inc.,: Machinery Fault Simulator - Lite* [online]. [visited on 2023-04-28]. Available from: https://spectraquest.com/machinery-fault-simulator/details/mfs-lt/.

65. SONG, Renwang; BAI, Xiaolu; ZHANG, Rui; JIA, You; PAN, Lihu; DONG, Zengshou. Bearing Fault Diagnosis Method Based on Multidomain Heterogeneous Information Entropy Fusion and Model Self-Optimisation. *Shock and Vibration* [online]. 2022, vol. 2022, e7214822 [visited on 2023-04-28]. ISSN 1070-9622. Available from DOI: 10.1155/2022/7214822. Publisher: Hindawi.

66. YUHONG, Jin; HOU, Lei; CHEN, Yushu. A new rotating machinery fault diagnosis method based on the Time Series Transformer. 2021.

67. MEY, Oliver; NEUDECK, Willi; SCHNEIDER, André; ENGE-ROSENBLATT, Olaf. *Machine Learning-Based Unbalance Detection of a Rotating Shaft Using Vibration Data* [online]. arXiv, 2020 [visited on 2022-10-31]. No. arXiv:2005.12742. Available from DOI: 10.48550/arXiv.2005.12742.

Appendix A: Resume

Appendix B: Plan of work

B.1 Summer semester - DP1

Period	Work
1st week	Consultation with the supervisor on directions of the future work based on literature review during previous semester.
2nd week	Outline the key sections of the analysis part in the thesis.
3rd week	Match supporting literature with analysis sections. Further invesigation on the feature engineering methodology in CbM.
4th week	Summarize notes from condition monitoring articles and video-recordings of tutorials and conferences.
5th week	Research transformation of vibration signal to feature space using time-frequency, harmonic and energy statistical metrics. Progress report meeting with the supervisor.
6th week	Find articles and take notes about unsupervised and semi-supervised techniques in streaming data for machinery diagnostics.
7th week	Narrow down wide variety applicable methods for signal decomposition.
8th week	Exploratory analysis on evaluation datasets. Progress report meeting with the supervisor to topic of related work.
9th week	Organize detailed outline out of notes gathered during literature research.
10th week	Write up the problem analysis about condition monitoring and evaluation datasets.
11th week	Write up the analysis section about feature engineering.
12th week	Write up the analysis section about machine learning diagnostics.

B.2 Winter semester - DP2

Period	**Work**
1st - 4th week	Apply feature engineering and classifications methods to evalution datasets
4th - 8th week	Design the vibration measurement for woodworking factoring according to technical standards.
8th - 12th week	Take preliminary measurements in the factory and compare signals with evaluation datasets.