

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

Návrh zadania diplomovej práce

Finálna verzia do diplomovej práce¹

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

Machinery vibrodiagnostics with the industrial internet of things

Miesto vypracovania: Ústav počítačového inžinierstva a aplikovanej informatiky,
FIIT STU, Bratislava

Oblast problematiky: Internet vecí, Spracovanie signálov, Feature engineering

Text návrhu zadania²

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 výchyiek operátorom a manuálna identifikácia súčastok 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 črt 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

Literatúra³

- 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 viest 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 / nie⁴

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

Declaration 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

Contents

1	Introduction	1
2	Problem analysis	3
2.1	Condition monitoring	3
2.1.1	Maintenance strategies	3
2.1.2	Vibration fault types	6
2.1.3	Technical standards	8
2.2	Signal preprocessing	12
2.2.1	Detrending	12
2.2.2	Adaptive noise cancellation	12
2.2.3	Time synchronous averaging	14
2.3	Feature engineering	14
2.3.1	Feature extraction	15
2.3.2	Time domain features	15
2.3.3	Frequency domain features	15
2.3.4	Time-frequency domain features	18
2.3.5	Feature transformation	26
2.3.6	Feature selection	27
2.4	Diagnostics techniques	29
2.4.1	Novelty detection	30
2.4.2	Classification	31
2.5	Evaluation datasets	33
2.5.1	Machinery Fault Database	33
2.5.2	CWRU bearings dataset	35
2.5.3	Unbalance of the rotating shaft	36

3 Design	39
3.1 Preparatory exploration	39
3.2 Vibration processing	40
Literature	43
A Plan of work	

List of Figures

2.1 P-F curve represents the evolution of the asset's health [7]	5
2.2 Bath tub curve [1]	6
2.3 Complex machinery vibrations [6]	6
2.4 The transducer linear response and resonance in tolerance intervals [11]	11
2.5 Transfer function of 1st order DC blocker filters [13]	13
2.6 Adaptive noise cancellation filter diagram	14
2.7 Comparison of time-frequency transform spectrograms [28]	21
2.8 Dyadic filter banks for discrete wavelet transform [16]	22
2.9 Empirical wavelet segmentation of Fourier spectrum [33]	24
2.10 Illustration of improved EWT spectral segmentations	25
2.11 DenStream [45]	30
2.12 Half-space tree [48]	31
2.13 Nearest neighbors classification algorithm [50]	32
2.14 Machinery fault simulator for MaFaulDa	34
2.15 CWRU machine apparatus	35
2.16 Motor driving shaft in unbalance measurement [59]	36
3.1 Machine learning lifecycle of feature discovery	40
3.2 Sensor network components	42

List of Equations

2.0	DC blocker IIR filter of 1st order	12
2.1	Adaptive noise filtering LMS update formula	13
2.2	Adaptive noise filtering NLMS update formula	13
2.3	Time synchronous averaging	14
2.4	MMS algorithm: Local maxima identification equality	17
2.6	MMS algorithm: Local minima identification equality	17
2.8	Harmonic series search criterion	17
2.9	Mother wavelet function	19
2.10	Continuous Wavelet transform	19
2.11	Representation of components for Synchrosqueezing Wavelet transform .	20
2.12	Teager–Kaiser energy operator	20
2.13	Transient-extracting transform	20
2.14	Wavelet packet coefficient	22
2.15	Empirical scaling function	23
2.16	Empirical wavelet function	23
2.17	Spectral negentropy	25
2.18	Min-max feature scaling	26
2.19	Feature standardization	26
2.20	Singular Value Decomposition	27
2.21	Pearson correlation coefficient	28
2.22	Fisher score	29
2.23	Mutal information	29

1 Introduction

Manufacturing is experiencing a shift in traditional 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 done on time in response to the key indicators.

The goal is to preserve required safety and production efficiency while extending the useful life of rotating mechanical 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 accurate 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 powered out of the battery pack. Preferably these devices should also remain physically small to reduce the additional clutter in the factory.

Attempted machine learning 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 during regular operation. The model has to be trained 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 quickly obtain a large enough sample of fault events. Novelty detection is a technique that can be applied in this case.

This progress report on the master's thesis is organized as follows. In the problem analysis, in section 2.1 we explore the mechanical maintenance approaches and industry standards on common fault identification. Section 2.2 summarizes filters for signal preprocessing. The 2.3 is about measuring vibrations and their transformation into features meaningful in automatic fault pattern recognition. In 2.4, we delve into modes of machinery diagnostics based on obtained relevant indicators, and 2.5 describes the most common datasets for machinery faults. Chapter 3 proposes the preliminary design of the model exploration and deployment phase of the project.

2 Problem analysis

In the problem analysis chapter we explore 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, inadequate 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. Vibration signals are the preferred tool for rotating machinery monitoring [1].

The defect needs to be 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 as 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 for spare parts to be ready in the supply inventory at all times. In certain situations, it suffices 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 enough. 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 when short downtime is acceptable, full and swift replacement of a broken machine with a backup is possible, or there is a negligible threat to the surrounding environment on failure [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 (MTTF - Mean Time To Failure). The schedule is crucial but can result in components being replaced in good condition creating waste. The parts can occasionally stay in operation too long, and the machine breaks as a result. Conservative planning is usually the norm to keep the machine always in a perfect state, and therefore more frequent interventions are required [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 the identification of troublesome components.

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].

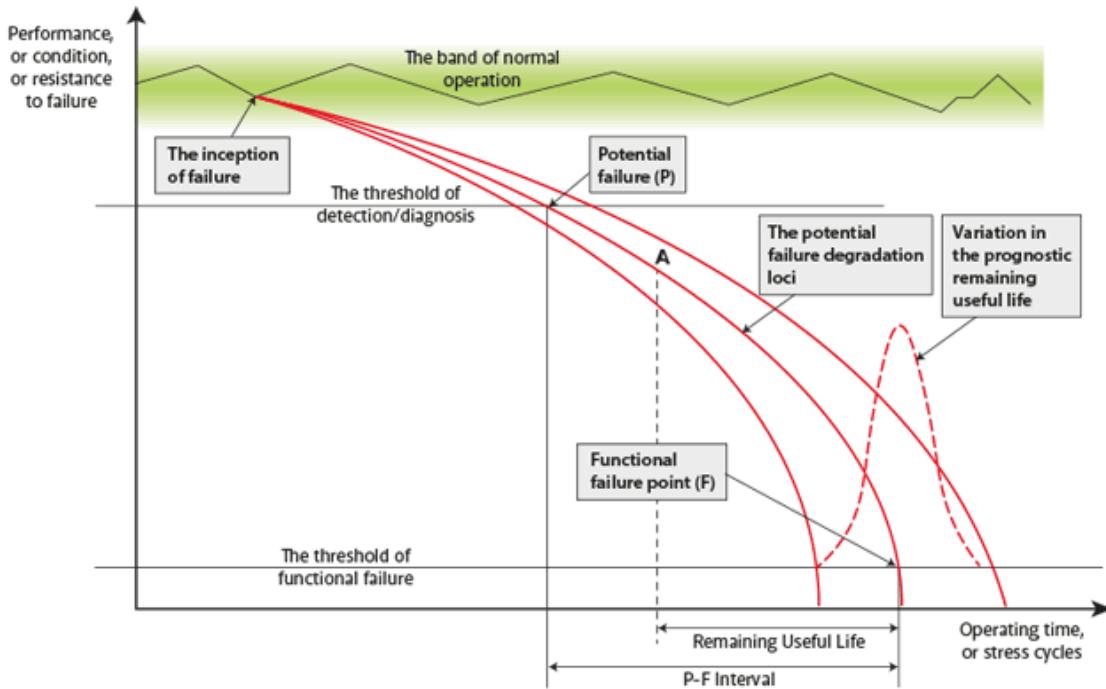


Figure 2.1: P-F curve represents the evolution of the asset's health [7]

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 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, based on the model of the ‘run-to-failure’ histories, and usage parameters [9]. Predictive condition monitoring aims to extend lifespan to the maximum by predicting expected RUL.

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. Manufacturing or material defects, inadequate installation, or improper start-up procedures, are all suspected causes. During the stable middle phase, malfunction can occur after the machine’s excessive overload. The time plot to failure rate is known as the bathtub curve (Fig. 2.2).

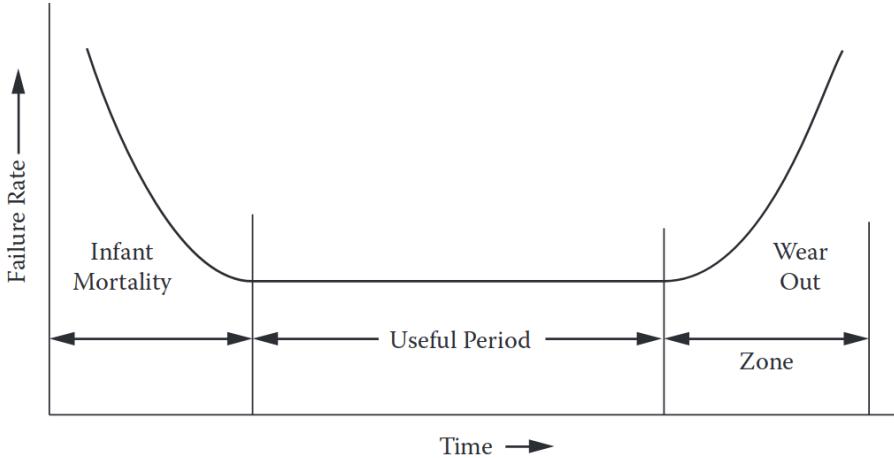


Figure 2.2: Bath tub curve [1]

2.1.2 Vibration fault types

Mechanical problems during machinery operation bring about vibrations in many instances. Therefore vibroacoustic diagnostics is considered as 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. These defects are broadly categorized as **unbalance, misalignment, looseness, eccentricity, deformation, crack, and influence of the external force** (friction) [6]. It is important to stress that our concern is not the underlying deformities in mechanical parts, but the correct fault classification based on the signal waveform.

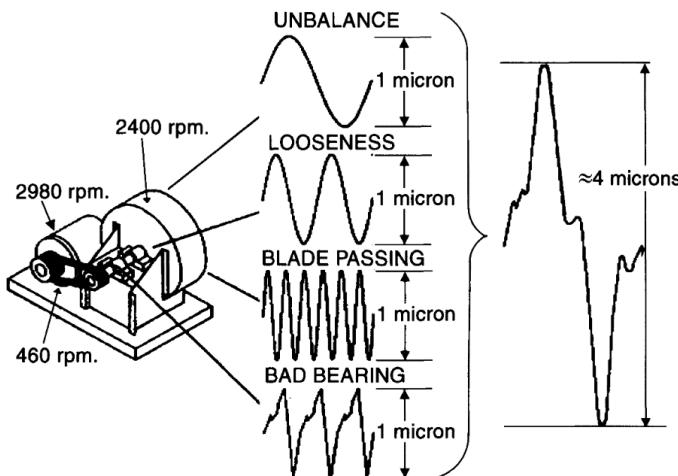


Figure 2.3: Complex machinery vibrations [6]

Rotating machine disorders exhibit frequency signatures at various ranges in the frequency spectrum with supplementary symptoms carried in phase signal. 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 and 1 kHz. Higher frequencies, measured traditionally to a limit of 10 kHz, help notice the faults 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**. The excessive peaks at harmonic frequencies are of interest. Harmonics 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**, **subynchronous**, or **non-synchronous components**. The defects can occur 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]. Procedures relying on elimination narrow 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. 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. Relevant standards for IoT diagnostics systems are *ISO 20816* (updated from ISO 10816) and *ISO 13373*.

ISO 20816-1:2016 establishes the approaches of 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 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 is related to signal energy. No simple direct relationship is expressible among these physical quantities, except in stationary signals.

The vibration severity is the maximum of magnitudes 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 different bearings or couplings, should be assessed independently.

Criteria introduced to judge vibration severity are 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 values in this table serve as guidelines towards realistic requirements between machine manufacturers and their customers.

Vibration velocity RMS [mm/s]	Class I Small machines	Class II Medium machines	Class III Large machines Rigid supports	Class IV Large machines Flexible support
0.28	Good (A)	Good (A)	Good (A)	Good (A)
0.45				
0.71				
1.12	Satisfactory (B)	Satisfactory (B)	Satisfactory (B)	Satisfactory (B)
1.8				
2.8	Unsatisfactory (C)	Unsatisfactory (C)	Unsatisfactory (C)	Satisfactory (B)
4.5				
7.1				
11.2	Unacceptable (D)	Unacceptable (D)	Unacceptable (D)	Unacceptable (D)
18				
28				
45				

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

Zone A is reserved for newly commissioned machines. *Zone B* signifies suitability for long-term operation. In *zone C* the machine is deemed in unsatisfactory condition and corrective action should be taken soon. Finally, in zone D vibrations can cause damage to the machine. The span of acceptable values differs with the machine class from I through to IV and their 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 providing 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 of 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 terms of the vocabulary in ISO 20816. According to the standard, the data collection operates in continuous or periodic observation modes which follow an event or interval. Both designs can be permanently mounted. In continuous collection it is recommended to have ‘multiplexing rate 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 program 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-monitoring 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 its identifier and type, power source, rated rotation speed and power, configuration (shaft or belt driven), and machine support. Measurement parameters are to be recorded alongside the measurement value itself 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.) [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 advise 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.

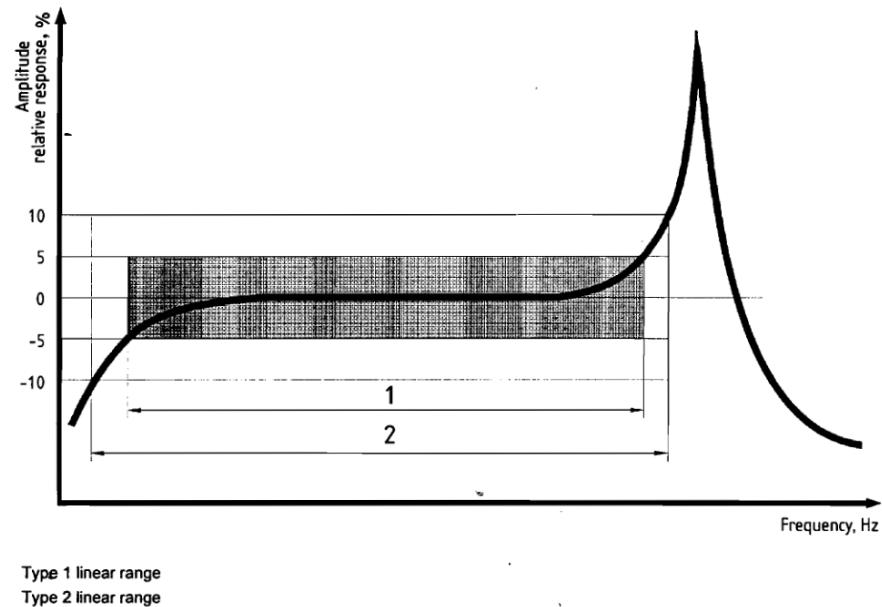


Figure 2.4: The transducer linear response and resonance in tolerance intervals [11]

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 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 the bathtub 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, the waveform in the 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 Signal preprocessing

The vibration signals in a factory environment are inherently full of disturbances. Nearby equipment operation and handling of heavy objects in the surroundings can all contribute to the unwanted chaotic movement in otherwise mostly pure oscillatory motion. In addition, accelerometers suffer from systematic measurement errors in the form of thermal noise, zero-g offset as a result of slight miscalibration, and bias originating from a constant force of gravity. These unavoidable distortions are suppressible to some extent with digital filters. In the preprocessing stage, we consider detrending, noise reduction with adaptive filters, and time synchronous averaging to eliminate the external interference.

2.2.1 Detrending

The oscillatory motion should be centered around the zero level for further manipulation. The constant offset is eliminated simply by subtracting the overall mean from the signal. Moreover high pass DC blocker infinite impulse response (IIR) filter of 1st order can adjust to shifts of the average value over time (Equation 2.1). The transition band depends upon the choice of corner frequency f_{3dB} (Fig. 2.5).

$$y_k = \left(1 - \frac{\omega}{2}\right) \cdot (x_k - x_{k-1}) + (1 - \omega) \cdot y_{k-1}; \quad \omega = 2\pi \cdot \frac{f_{3dB}}{f_s} \quad (2.1)$$

A steeper 3 dB attenuation band can be achieved by increasing the order of the filter. Then the cutoff frequency should be such that filter coefficients are fractions to counteract rounding errors [13].

The finite impulse response (FIR) filter is not recommended for DC component removal because of the undesirable ripple effect with the small number of taps. Cascaded-integrator-comb (CIC) filters are proposed as an alternative instead [14].

2.2.2 Adaptive noise cancellation

Adaptive noise cancellation (ANC) involves an adaptive filter that self-adjusts coefficients through an update algorithm in response to the reference noise signal. The objective of this filter is to minimize the mean square error (MSE) cost function in

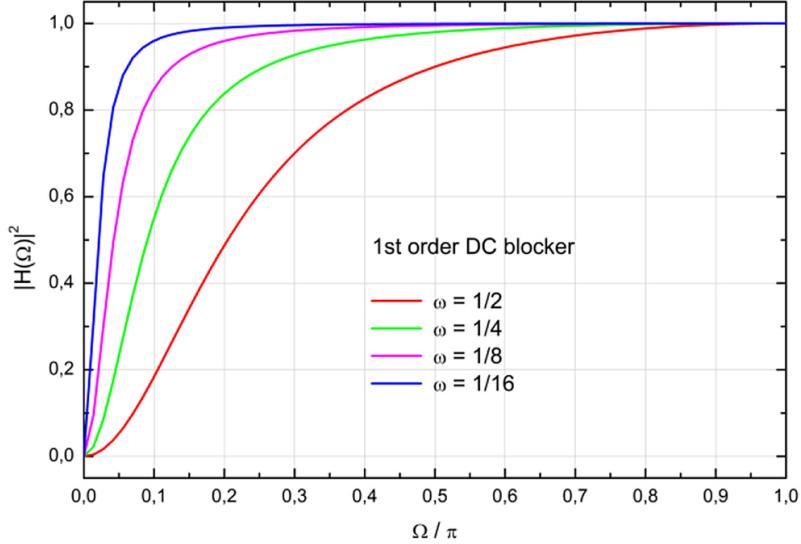


Figure 2.5: Transfer function of 1st order DC blocker filters [13]

the error signal e_k between signal contaminated with Gaussian noise d_k and filter output y_k . Additive noise n_k is assumed to be correlated with noise signal \mathbf{X}_k [15].

Wiener-Hopf equations solve the optimal gradient of the MSE function and provide us with FIR filter coefficient vector \mathbf{W}_k . The least mean squares (LMS) algorithm recursively approximates this analytical solution with the method of steepest descent (Equation 2.2) [15]. The multiple parameters are to be considered in the evaluation of filter performance: convergence rate, estimated error, and signal-to-noise ratio (SNR).

$$\mathbf{W}_{k+1} = \mathbf{W}_k + 2\mu \mathbf{X}_k e_k \quad (2.2)$$

The convergence stability is affected by step size μ which is bounded from above with the inverse of the maximal eigenvalue of input covariance matrix λ_{max} . The normalized least mean squares (NLMS) can handle input of varying scales (Equation 2.3).

$$\mathbf{W}_{k+1} = \mathbf{W}_k + \frac{\mu}{\|\mathbf{X}_k\|^2} \mathbf{X}_k e_k \quad (2.3)$$

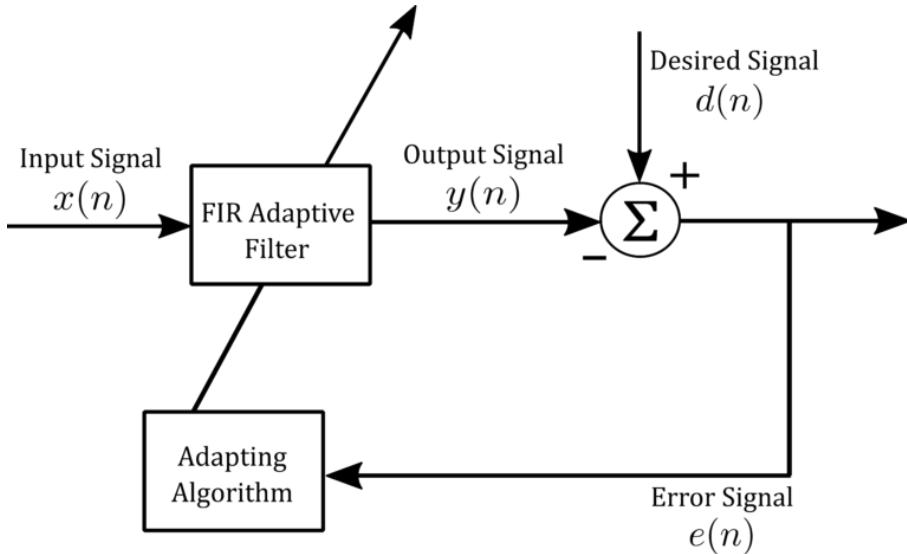


Figure 2.6: Adaptive noise cancellation filter diagram

2.2.3 Time synchronous averaging

Time synchronous averaging (TSA) diminishes the impact of vibration sources unrelated to the rotational frequency and its harmonics. TSA averages time-domain waveform over N points and aligns it to a synchronization pulse with period T (Equation 2.4). This technique has been successfully applied to the gearbox and bearing fault diagnosis [6, 16]

$$x_{TSA} = \frac{1}{N} \sum_{n=0}^{N-1} x(t + nT) \quad (2.4)$$

2.3 Feature engineering

Raw numerical vectors after preprocessing are merely low-level descriptors of the underlying physical phenomena. At first, these incomprehensible sequences of numbers are reduced to summary attributes called features in the process of feature discovery. Normalization and linear transformations are applied here to discriminate different categories in the feature space. The most informative set of features is obtained with feature selection methods for the diagnostic model. Features can be hand-crafted, learned implicitly within the model representation, or explicitly from an optimization problem solution.

2.3.1 Feature extraction

Predictive maintenance has ideal prerequisites for the application of feature engineering because the signal is usually pseudo-stationary, and the trend monitoring variables come out of extensive domain expertise in mechanics. The advantages of add-in extraction effort, as opposed to processing samples unmodified, are to gain better classification precision, reduce computational burden and storage capacity downstream with dimensionality reduction [17]. It is important to note that the design of features is not a standalone step in the machine learning pipeline but it should be performed iteratively to improve the target model. Signal features are computed in the time, frequency, and time-frequency domain [18].

2.3.2 Time domain features

The most widely found features in the literature are rudimentary statistical measures of the central moment: mean, variance, standard deviation, skewness, and kurtosis (Tab. 2.3). Statistics can be calculated in any domain, but the mean value is not to be used in detrended data. The vibration severity metrics out of technical standards are also highly regarded. The characteristics of amplitude include root mean square, peak-to-peak, and maximum [19].

The other significant time-domain attributes are derived as ratios of previous simpler ones. These ratios are crest factor, margin factor, impulse factor, and shape factor (Tab. 2.3) [16]. Many articles have been successful in bearing fault detection out of transients in impulsive signals with kurtosis, crest factor, and margin indicators [18]. It is also suggested that the shape factor can signify unbalance and misalignment faults [16].

2.3.3 Frequency domain features

The mechanical faults present themselves as oscillatory patterns which are combinations of frequencies with various amplitudes. The Fourier transform is one of the prominent strategies in power spectral density estimation. Experts on vibrodiagnostics utilize it as a primary signal processing technique for data analysis as it is

Feature	Equation	Feature	Equation
Standard deviation	$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2}$	Crest factor	$X_{cf} = \frac{\max(x_i)}{\left(\frac{1}{N} \sum_{i=1}^N x_i^2\right)^{\frac{1}{2}}}$
Skewness	$X_{sv} = \frac{1}{N} \sum_{i=1}^N \left(\frac{x_i - \bar{x}}{\sigma}\right)^3$	Margin factor	$X_{mf} = \frac{\max(x_i)}{\left(\frac{1}{N} \sum_{i=1}^N \sqrt{ x_i }\right)^2}$
Kurtosis	$X_{kv} = \frac{1}{N} \sum_{i=1}^N \left(\frac{x_i - \bar{x}}{\sigma}\right)^4$	Impulse factor	$X_{if} = \frac{\max(x_i)}{\frac{1}{N} \sum_{i=1}^N x_i }$
Root mean square	$X_{rms} = \left(\frac{1}{N} \sum_{i=1}^N x_i^2\right)^{\frac{1}{2}}$	Shape factor	$X_{sf} = \frac{\left(\frac{1}{N} \sum_{i=1}^N x_i^2\right)^{\frac{1}{2}}}{\frac{1}{N} \sum_{i=1}^N x_i }$
Peak-to-peak	$X_{ppv} = \max(x_i) - \min(x_i)$	Maximum	$X_{max} = \max(x_i)$

Table 2.3: Time domain features

recommended in ISO 13373-2 standard [20].

The inherent symmetries in the Fourier matrix made it possible to invent an efficient implementation of the Fast Fourier transform (FFT) algorithm with time complexity $O(n \log n)$. The drawback of the plain spectral analysis is the lack of resolution for events that occurred at distant time instants, and so their spectral components might adversely blend in together.

In the frequency domain, we can obtain spectral centroid, spectral kurtosis, spectral roll-off, spectral flux, energy in frequency bands, and energy ratio (Tab. 2.4) [21]. In geometry terms, the spectral centroid represents the barycenter of the frequency magnitude plot. Spectral roll-off gives a notion about the spectral distribution because it searches the frequency f_c below which 95% of the signal energy is contained. The energy in the roll-off calculation is summed up to the Nyquist limit: $f_s/2$. According to the definition, spectral flux is normalized cross-correlation between two successive amplitude spectra, value of one means spectra are the most dissimilar.

The harmonic frequencies may originate from higher order oscillation modes of the same shaft or other elastic rod supported on both ends. Several harmonics features worth investigating are fundamental frequencies, noisiness, inharmonicity, and harmonic spectral deviation (Tab. 2.4). Noisiness as a harmonic feature is a ratio of noise (non-harmonic components) energy to the total contained energy. Inharmonicity measures the divergence of spectral components from a purely harmonic to inharmonic signal on scale from 0 to 1. Harmonic spectral deviation adds up differences of amplitudes at harmonic peaks $a(h)$ from the spectral envelope $SE(h)$ [21].

If a single principal frequency exists, it can be determined with maximum likelihood estimation. Such frequency would explain the signal spectrum the best [21]. The frequency spectrum is a discrete set of amplitudes where peaks have to be reliably identified to create representative attributes. The essential peak-finding approaches are based either on magnitude or gradient. All found extrema are commonly filtered with the magnitude of prominences and the widths at half prominence. In the magnitude-based method, middle point x_i is compared to neighboring two points and the peak is then: $x_{i-1} < x_i > x_{i+1}$. The gradient-based method evaluates the first derivative at the point which is equal to zero in case the point is a local maximum, local minimum, or inflection point [22].

A substantial improvement is a robust non-parametric peak identification named *MMS* based on the sum of terms in an arithmetic progression based on maximum, minimum, and sum. MMS max-min finder in the elementary form processes points in the window of length 3, it advances one point and deems its middle point as a local extremum if it satisfies equalities below. Equation 2.6 is for the hill and Equation 2.8 is for the valley. The filtration techniques are incorporated in the adaptations of the MMS algorithm: MMS-WBF, MMS-SG, MMS-LH [22].

$$\text{MMS}_{\max} = \text{MMS}_{\max|\text{mid}} \quad (2.5)$$

$$\frac{a_{\max} - a_{\min}}{S_3 - a_{\min} \cdot 3} = \frac{a_{\text{mid}} - a_{\min}}{S_3 - a_{\min} \cdot 3} \quad (2.6)$$

$$\text{MMS}_{\min} = \text{MMS}_{\min|\text{mid}} \quad (2.7)$$

$$\frac{a_{\max} - a_{\min}}{a_{\max} \cdot 3 - S_3} = \frac{a_{\max} - a_{\text{mid}}}{a_{\max} \cdot 3 - S_3} \quad (2.8)$$

Multiple harmonic series and the sidebands can be separated into a discrete set of frequency components, each with central frequency, uncertainty, and amplitude $C_i(v_i, \Delta v_i, A_i)$ by an exhaustive search algorithm. Harmonic family identification is a non-trivial problem because of the spectrum estimation errors. The criterion is proposed to select harmonic at the minimal distance from the true fundamental frequency multiple (Equation 2.9). Two series with the same fundamental frequency are merged and thought of as a modulation series [23].

$$v_i^{(r)} = \frac{v_j}{\min |v_j - r \cdot v_i|} \quad (2.9)$$

Feature	Equation
<i>Spectral centroid</i>	$X_{fc} = \frac{\sum_{i=0}^{N-1} f_i \cdot a(f_i)}{\sum_{i=0}^{N-1} a(f_i)}$
<i>Spectral roll-off</i>	$\sum_0^{f_c} a^2(f) = 0.95 \sum_0^{f_s/2} a^2(f)$
<i>Spectral flux</i>	$X_{flux} = 1 - \frac{\sum_k a(t-1,k) \cdot a(t,k)}{\sqrt{\sum_k a(t-1,k)^2} \sqrt{\sum_k a(t,k)^2}}$
<i>Harmonic spectral deviation</i>	$X_{HDEV} = \frac{1}{H} \sum_h (a(h) - SE(h))$
<i>Noisiness</i>	$X_{noise} = \frac{\sum_{k \notin h}^N E_k}{\sum_{i=1}^N E_i}$
<i>Inharmonicity</i>	$X_{inharmo} = \frac{2}{f_0} \cdot \frac{\sum_h f(h) - h \cdot f_0 \cdot a^2(h)}{\sum_h a^2(h)}$
<i>Energy</i>	$E = \sum_{i=0}^{N-1} x(t) ^2$
<i>Energy ratio</i>	$E_i = \frac{E_i}{\sum_{i=1}^N E_i}$
<i>Shannon entropy</i>	$H(X) = - \sum_i P(X = x_i) \cdot \ln(P(X = x_i))$

Table 2.4: Frequency domain features

2.3.4 Time-frequency domain features

The **Short-time Fourier transform** (STFT) splits the time-domain signal to an equal-length intervals. Individual chunks have 50% overlap and are multiplied with weights of window function to balance scalloping loss and spectral leakage due to Fourier transform periodicity assumption. Traditionally, the Hann window is commonly used instead of rectangular window [5, 20].

In the time-frequency domain, the same features can be derived as in the frequency domain, but in addition, the attributes are time-localized in this way. The STFT has a considerable flaw for implementation in a self-adaptable system and that is fixed resolution. The optimal window size has to be set beforehand or chosen after performing multiple transformations on chunks out of the range of lengths. Welch's methods averages multiple consecutive blocks to better estimate the spec-

trum. We have already researched the suitability of STFT for online detection of constant frequencies [24].

The bands for lower frequencies should be longer in duration than for higher frequencies. The **Wavelet transform** (WT) possesses such multi-scale discrimination property effectively increasing resolution in time-frequency plain. Wavelet basis functions are constructed for that purpose (Equation 2.10). There are several wavelet families, for example, Haar, Daubechies, Coiflets, Symlets, Morlet, and Meyer [16].

$$\psi_{s,\tau} = \frac{1}{\sqrt{s}}\psi\left(\frac{t-\tau}{s}\right) \quad (2.10)$$

Continuous Wavelet transform (CWT) (Equation 2.11) is performed by scaling and translating the mother wavelet ψ picked out of the appropriate family [16]. The scale factor is denoted with s and time position with τ . The choice of wavelet type is data-driven because distinct wavelet shapes have an impact on the response and ultimately contribute to filter length. The decision lies between recognition abilities for impulse-like signals or the inclusion of wider surrounding space.

$$W_{x(t)}(s, \tau) = \frac{1}{\sqrt{s}} \int x(t) \cdot \psi^*\left(\frac{t-\tau}{s}\right) dt \quad (2.11)$$

The CWT is computationally intensive when a highly detailed scale resolution is required because each wavelet scale convolves with the entire signal. The fCWT algorithm allows 100 times higher spectral resolution than previous implementation at the same speed. It increases performance 122 times compared to Wavelib and 34 times in comparison with PyWavelets [25].

The fast CWT algorithm reaches compelling improvement by applying Parseval's theorem to the wavelet transform formula that removes the dependence on the time offset parameter [25]. The convolution takes place with the mother wavelet in the Fourier base. Then, inverse FFT produces the coefficients for individual scales.

Synchrosqueezing Wavelet transform (SST) is a modification of CWT attempting to sharpen the representation of frequency components by coefficient reassignments from around the central frequencies towards the middle of the bands. The justification for these reallocations is rooted in the signal approximation as

amplitude-modulated oscillating modes with additive noise $\eta(t)$ (Equation 2.12) [26].

$$s(t) = \sum_{k=1}^K A_k(t) \cos(\theta_k(t)) + \eta(t) \quad (2.12)$$

The components are defined by their instantaneous amplitudes $A_k(t)$ and instantaneous phases $\theta_k(t)$. The energy spread to adjacent bins can be effectively squeezed only in regions with constant phase and large enough component separation. Despite the promising properties of this transform, white noise causes severe interference in the resulting time-frequency map.

The alternative to isolating weak impacts with high time resolution is a **Teager-Kaiser energy operator** (TKEO) (Equation 2.13). It is a tool for envelope analysis to demodulate characteristic AM-FM signal present during bearing faults. Energy operator output can be utilized as a standalone feature attribute.

$$\psi[x(n)] = [x(n)]^2 - x(n-1)x(n+1) \quad (2.13)$$

Improved TKEO is necessary to prevent analysis from suffering from the noisy source. The key idea is to perform TKEO after signal decomposition into narrow-band components with different center frequencies. The extracted modes are reconstructed with weights assigned based on their correlations to the original signal [27].

Time-frequency spectrum modification preserving localization of abrupt wide-band spikes at time t_0 and simultaneously reducing energy smear over the larger region is a goal of **Transient-extracting transform** (TET). Post-processing of STFT window $G(t, \omega)$ involves multiplication of the spectrum with the Transient extracting operator (TEO). This operator is expressed in the form of a Dirac delta function $\delta(t)$ (Equation 2.14).

$$\text{Te}(t, \omega) = G(t, \omega) \cdot \delta(t - t_0) \quad (2.14)$$

TET representation retains non-zero coefficients where the absolute value of the ratio between two STFTs $G^{tg}[n, k] / G[n, k]$ is less than half the sampling interval T . These two transforms use distinct windows $g[n]$ and $n \cdot g[n]$. Decomposition of signal with TET is proved to produce significantly larger kurtosis (around 38 in TET, 4 in other methods) and hence better discriminate the transient fault [28].

The spectrograms to illustrate the difference in the ability of Fourier transform, continuous Wavelet transform, and their modifications to pick up underlying patterns in bearing faults are shown in the Fig. 2.7.

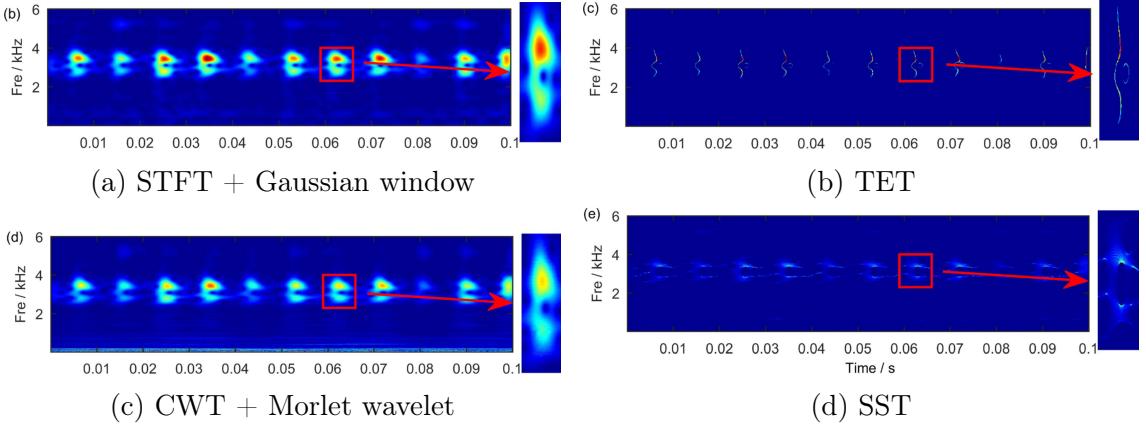


Figure 2.7: Comparison of time-frequency transform spectrograms [28]

The dyadic filter bank is another signal decomposition technique that generates subbands on multiple granularity levels. The practical realization of the multi-scale description is **Discrete Wavelet Transform** (DWT). The DWT behaves as a quadrature mirror filter and splits waveform using a wavelet filter to detail coefficients (D_1) and approximation coefficients (A_1) (Fig. 2.8a) [16]. Low-pass filter $h(k)$ creates approximation coefficients further decomposed in the successive levels. Detail coefficients represent the result of the high-pass filter $g(k) = (-1)^k h(1 - k)$ after decimation by the factor of 2.

The maximum depth of the decomposition tree is $\log_2 n$ where n is the number of input samples. Energy, energy ratio, and entropy are prevalent features that succinctly encode the wavelet coefficients. Otherwise, the additional extracted levels raise the total number of data dimensions.

In washing machine status classification, the discrete wavelet transform with Daubechies wavelet (db4) and fifth-level decomposition provided features combined from approximation (cA_5) and detail coefficients (cD_1, \dots, cD_5). Washing machines belonged to three categories: no fault, electric motor clamping screws problem, and a loose or broken counterweight. Extracted measures were sample mean and sample variances over autocorrelation functions of coefficients (AcD_n) and smoothed coefficients cD_1, cD_2 by moving average filter [29].

Wavelet Packet Decomposition (WPD) applies filters to split detail coefficients identically as approximation ones (Fig. 2.8b) thus increasing the resolution in the high-frequency bands and providing uniform spectrum partitioning.

$$w_{j,n,k} = \langle f, W_{j,k}^n \rangle = \langle f, 2^{j/2} W^n(2^j t - k) \rangle \quad (2.15)$$

Each wavelet packet coefficient $w_{j,n,k}$ captures subband frequency content around time instant $2^j k$ (Equation 2.15) [30]. This measure is an inner product of the source and scaled wavelet packet function. The aforementioned feature extraction established in DWT can be applied, for example calculation of the wavelet packet node energy.

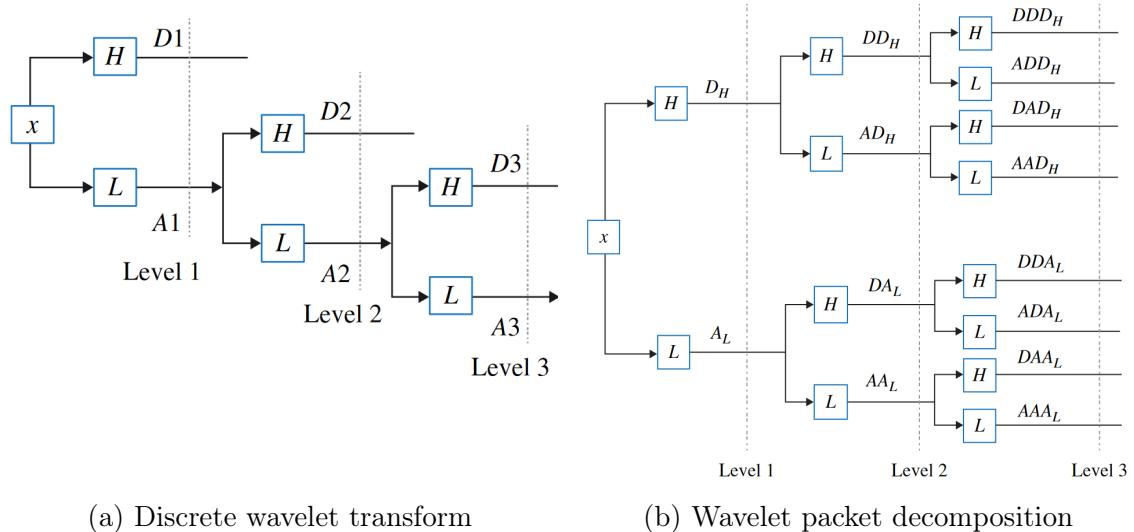


Figure 2.8: Dyadic filter banks for discrete wavelet transform [16]

The wavelet packet energy ratio for weak feature extraction has been incorporated into the method of multiple frequency bands demodulation (MFBD). The highest n energy coefficients are selected out of 8 narrow frequency bands to subsequently affect the principal components. Demodulation principle for the frequency bandwidth prescribes n to satisfy condition: $1 / 2^n > f_{\text{modulation}} / f_s$. The first few eigenvectors explaining together more than 80% of energy are retained to reconstruct the signal with Fourier transform and retain the weak fault components [31].

Tool wear diagnosis based on acoustic emission (AE) signal considers *wavelet packet energy* in bands E_8 , E_{10} , E_{12} , *energy ratios* P_8 , P_{13} , and *energy entropy* as having high correlation ($|r| > 0.8$) with the band saw flank face width. The

acoustic signal is decomposed into three layers using Daubechies db3 wavelet. The bottom layer contains bands numbered 7 through 14, each with a bandwidth of 62.5 kHz, because of the 1 MHz sampling frequency. The feature vector constructed in the article includes other statistical metrics out of power spectral density that has reached a notable correlation with evolving wear. The statistics are *skewness*, *kurtosis*, *shape factor*, and *centroid frequency* [32].

Discrete wavelet transform and wavelet packets partition the spectrum into pre-defined frequency bands that do not always adequately capture individual elementary oscillations. Adaptive spectral segmentation is needed to extract separate intrinsic mode functions (IMF).

Empirical Wavelet Transform (EWT) constructs adaptive bandpass filters with Meyer wavelet. The scaling function $\hat{\phi}_n$ and wavelet function are defined according to equations (2.16) and (2.17) respectively. The inner product of the signal with the scaling function obtains the approximation coefficients and the inner product with the wavelet function results in detail coefficients.

$$\hat{\phi}_n(\omega) = \begin{cases} 1 & \text{if } |\omega| \leq \omega_n - \tau_n \\ \cos \left[\frac{\pi}{2} \cdot \beta \left(\frac{1}{2\tau_n} (|\omega| - \omega_n + \tau_n) \right) \right] & \text{if } \omega_n - \tau_n \leq |\omega| \leq \omega_n + \tau_n \\ 0 & \text{otherwise} \end{cases} \quad (2.16)$$

$$\hat{\psi}_n(\omega) = \begin{cases} 1 & \text{if } \omega_n + \tau_n \leq |\omega| \leq \omega_{n+1} - \tau_{n+1} \\ \cos \left[\frac{\pi}{2} \cdot \beta \left(\frac{1}{2\tau_{n+1}} (|\omega| - \omega_{n+1} + \tau_{n+1}) \right) \right] & \text{if } \omega_{n+1} - \tau_{n+1} \leq |\omega| \leq \omega_{n+1} + \tau_{n+1} \\ \sin \left[\frac{\pi}{2} \cdot \beta \left(\frac{1}{2\tau_n} (|\omega| - \omega_n + \tau_n) \right) \right] & \text{if } \omega_n - \tau_n \leq |\omega| \leq \omega_n + \tau_n \\ 0 & \text{otherwise} \end{cases} \quad (2.17)$$

The normalized frequency axis in range $\omega \in [0, \pi]$ is divided by split points $\omega_0, \dots, \omega_N$ where $\omega_n = f_n \cdot 2\pi / f_s$ (Fig. 2.9). Each segment is bounded between $[\omega_{n-1}, \omega_n]$ with transition phase of width $2\tau_n$ and polynomial transition function $\beta(x)$. A tight frame set of empirical wavelets is built by setting the transition phase

proportional to the band boundary: $\tau_n = \lambda \omega_n$, and proportional constant must obey constrain: $\lambda = \min \left(\frac{\omega_{n+1} - \omega_n}{\omega_{n+1} + \omega_n} \right)$. The N boundaries defining different portions of the Fourier spectrum are placed at the center between two consecutive local maxima [33].

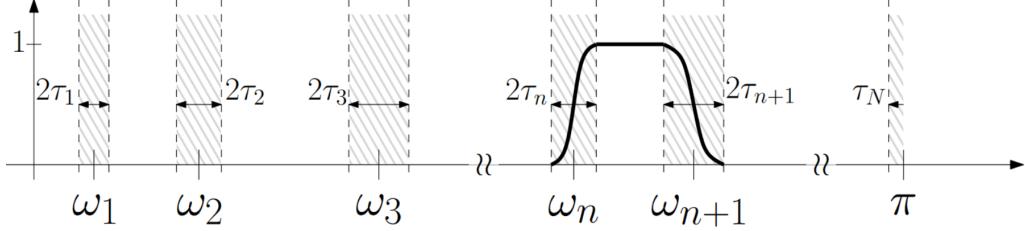


Figure 2.9: Empirical wavelet segmentation of Fourier spectrum [33]

The drawback of EWT is improper segmentation in noisy and non-stationary signals producing too many uninformative partitions of the spectrum. The solutions we discuss fall into two groups that apply improvements before or after the wavelet transformation step.

Refinements acting in the onset as preparation before wavelet processing create an envelope curve in the frequency domain and attempts to eliminate supposed interference.

A combination of **Maximum Correlated Kurtosis Deconvolution** and improved EWT (MKCD-EWT) favors periodic impacts by dynamically constructing an optimal FIR filter that maximizes the correlation kurtosis of the signal. The envelope curve smooths of the amplitude spectrum by linear interpolation (Fig. 2.10a). The threshold accounting for desired SNR modifies the envelope when more than n components are discovered (A_h is maximal amplitude and A_l is minimal amplitude): $\lambda = A_l + \frac{C}{\text{SNR}} (A_h - A_l)$. *Squared envelope spectrum* and *Teager energy operator spectrum* are computed as fault features for the IMF with highest kurtosis [34].

The segmentation boundaries detection can be accomplished also on the spectral envelope consisting of **Piecewise Cubic Hermite Interpolating Polynomial** (PCHIP-EWT) instead of scanning the Fourier spectrum directly (Fig. 2.10b). Divisions of the upper cut-off frequency are kept if local power indicates a subband with useful information ($p_i(f) \geq \lambda$) as in the case of abrupt level change. Local power is the ratio of n th local envelope maximum and difference of indexes of its adjacent local minima: $p_i(f) = K_{\max(i)(f)} / (f_{k_{\min(i+1)}} - f_{k_{\min(i)}})$. The best λ is set

experimentally [35].

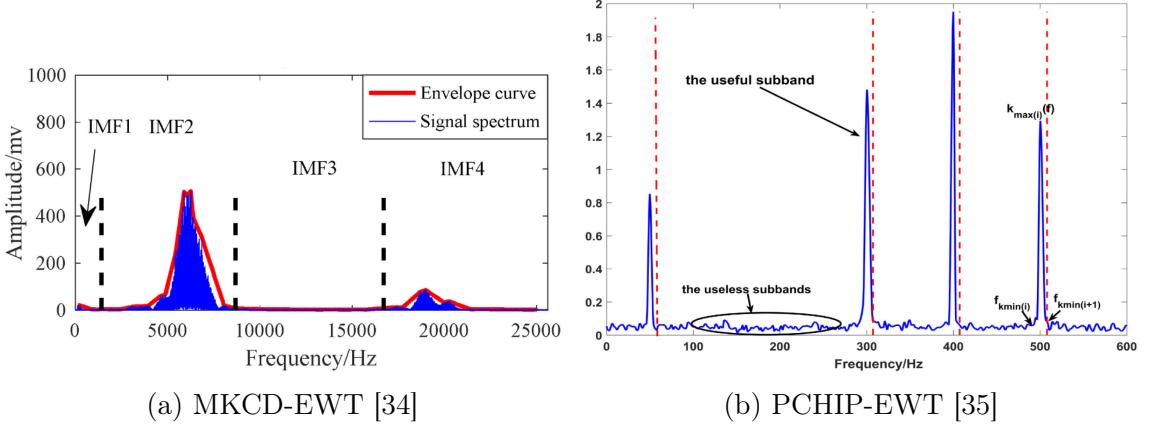


Figure 2.10: Illustration of improved EWT spectral segmentations

Alternatively, the central frequencies of bands containing the most unevenness measured by negentropy are picked after the EWT processing passed. The filter bank is then reconstructed accordingly. ‘Negentropy measures the inclination of a system to increase its level of organization’ [36]. The larger *spectral negentropy* (Equation 2.18) suggests more fault-induced impulses. $E(k, f, \Delta f)$ denotes the squared envelope spectrum (SES) which is an envelope in the Fourier domain: $\mathcal{F}\{|y(k, f, \Delta f)|^2\}$.

$$\Delta I_E(f, \Delta f) = \sum_{k=0}^{N-1} \frac{E(k, f, \Delta f)^2}{\frac{1}{N} \sum_{k=0}^{N-1} E(k, f, \Delta f)^2} \cdot \ln \left(\frac{E(k, f, \Delta f)^2}{\frac{1}{N} \sum_{k=0}^{N-1} E(k, f, \Delta f)^2} \right) \quad (2.18)$$

In **time-frequency domain scanning empirical spectral negentropy** method (T-FSESNE) the scanning EWT filter establishes the empirical mode components $X_i(t)$. The central frequencies f_{cj} are retained in resonance bands where the frequency-domain spectrum negentropy (FSNE) is greater than mean $\overline{\Delta I_E}$. The optimal bandwidth for each component is searched among the set of incrementally expanding regions fixed at f_{cj} . The interval with the largest time-domain spectrum negentropy (TSNE) is selected to restrict the bounds of the mode (f_{cj}, B_{wk}) [37].

The following scheme termed **adaptive and fast empirical wavelet transform** (AFEWT) calculates spectral negentropy of key functions $K'_i(f)$. The spectral section with the highest entropy is assumed to contribute the most to the rolling

bearing fault. First, the Fourier transform of the spectrum establishes the key function $K(f)$. This function is analogous to the cepstrum. The leftmost ridge in the $K(f)$ plot is transformed using inverse FFT to attain trend spectrum $T(f)$. Finally, the array of local minima points in the trend spectrum bound frequency bands whose Fourier transform produces key functions $K'_i(f)$ [38].

The inspiration for filtering with adaptive basis comes from **Empirical mode decomposition** (EMD). This procedure locates local maxima and minima in the time waveform that are interpolated with cubic splines to form signal envelopes. The average of the upper and lower envelope is repeatedly subtracted from the residuals recovering the higher order and lower frequency IMFs [39].

The EMD is not recommended for practical applications because it suffers from mode mixing problems and lacks rigid theoretical foundations. Many improvements based on the mode sifting process, such as Local Mean Decomposition (LMD), Ensemble Empirical Mode Decomposition (EEMD), Concealed Component Decomposition (CCD) [40], and Empirical Wavelet Transform, have been devised as being capable of estimating more reasonable modes.

The time-frequency representation offers plenty of features depending on the choice of spectral transformation. We have reviewed Short-time Fourier transform, Continuous Wavelet transform, Synchrosqueezing transform, Discrete Wavelet transform, Wavelet packet decomposition, Empirical Wavelet transform, and operators of negentropy, TKEO, and TEO.

2.3.5 Feature transformation

Numeric features from the feature extraction phase have non-normal distributions and span the range of scales. Broad differences among features skew the spread in the particular axis. Inevitably it can degrade the discernment of fault diagnostics models that map input onto smooth function as regression does [41]. The feature scaling, power transform, and principal component analysis modify attribute values to gain more meaningful predictors, but one must be cautious in model interpretation.

$$\tilde{x} = \frac{x - \min(x)}{\max(x) - \min(x)} \quad (2.19)$$

$$\tilde{x} = \frac{x - \bar{x}}{\sigma_x} \quad (2.20)$$

Feature normalization most often takes two forms. **Min-max scaling** changes original range of values into interval $[0, 1]$ (Equation 2.19). **Standardization** (Equation 2.20) constrains the mean of the variable to 0 with a variance of 1 [41].

Highly correlated features generated from a relatively small original column space are redundant as they do not provide any additional information for diagnostics. **Principal Component Analysis** (PCA) solves this problem by projecting potentially linearly dependant features into a new feature space where the incoming information is preserved in a smaller number of features [41]. The threshold of how many principal components are picked depends on the amount of explained variance and desired quantity of data reduction.

$$\mathbf{C} = \mathbf{U}\Sigma\mathbf{V}^T \quad (2.21)$$

PCA consists of taking the Singular Value Decomposition (SVD) (Equation 2.21) of the mean-centered input matrix. The disadvantage of this method is the loss of explainability in transformed space though it generally outperforms the model working with hand-crafted features [18]. The signal samples can be processed directly by PCA without going through an intermediate step of calculating statistical measures.

2.3.6 Feature selection

Not all features contribute to the model's discriminatory power with an even share. A certain subset can reach better results than other substitute options. The *filtering methods* of feature selection rank the predictors in order of their importance for the problem at hand, or they separate the group of predictors that achieve the top modeling accuracies in unison [17].

Still, the features can be chosen by *embedded methods* intrinsically as a part of a model or else in *wrapper methods* by machine learning search algorithm at a serious computational expense. However, we will focus next only on filtering methods.

The general steps in selecting the adequate predictors is outlined next [16]:

1. **Subset generation** - sets of features are generated in different search direc-

tions and with various strategies. Attributes are either appended to an empty set or pruned away from a universal set, sequentially or randomly.

2. **Subset evaluation** - comparison of subset quality is assessed with relevance measure some of which are discussed below.
3. **Stopping criteria** - search is exhausted when the specified number of features has been found, subset metrics cannot be improved further, or satisfactory model performance is achieved. Subset generation and evaluation can be performed multiple times until the stopping criteria are met.
4. **Validation** - resulting subset is tested for the specific model on synthetic and real-world datasets against well-known results.

Filter-based feature selection is preprocessing step independent of model choice with small computational requirements. Measures of information, correlation, similarity, and interdependence output the relevancy rating. Predictors are rated individually or in interacting congregations.

Most of the scores are based on supervised learning, so they expect true class labels to apportion the measurements respectively. After the scores are assigned to the first n features, those below a threshold are removed.

The frequently used scores upon which the feature relevance is ordered are [16]:

- **Variance threshold** - removes low-variance features below set threshold.
- **Pearson correlation** - expresses linear relationship between class label c and features f_i . Attributes are ranked in descending order according to the absolute value of their correlation coefficient from zero upwards, i. e. we seek the highest correlations to the class label.

$$r(i) = \frac{\text{cov}(f_i, c)}{\sqrt{\text{var}(f_i) \cdot \text{var}(c)}} \quad (2.22)$$

- **Fisher score** - measures the difference among the means of the classes. It is interchangeable with ANOVA F-value, but it is evaluated for each feature X^j separately. Ideally, the features in the subset have large distances between samples of various classes in C and distances within a class are the smallest

possible. In the formula (2.23), n_j is the sample size of j th feature, μ^j is its sample mean, and μ is the overall mean.

$$\text{FS}(X^j) = \frac{\sum_{i=1}^C n_i (\mu_i^j - \mu)^2}{\sum_{i=1}^C n_i (\sigma_i^j)^2} \quad (2.23)$$

- **Mutal information** - quantifies the dependence among features, or between features and class labels. It is almost identical to Information Gain. The probability distribution proximity of variables derives from relative entropy known as the Kullback-Leibler distance. Probabilities $P(x)$, $P(y)$, $P(x, y)$ are estimated in the contingency table from event occurrence count to all sample population $|x| / N$. Joint probability $P(x, y)$ represents samples of feature x simultaneously in class y .

$$\text{MI}(X, Y) = \sum_{y \in Y} \sum_{x \in X} P(x, y) \log \left(\frac{P(x, y)}{P(x)P(y)} \right) \quad (2.24)$$

Multiple subsets of predictors produced by each evaluation metric can train several variants of a classification model. Sets of attributes can be likewise combined into an ensemble by majority voting, taking the best features out of every group.

2.4 Diagnostics techniques

Fault identification in the rotating machinery is one-class or multi-class classification problem acting in a semi-supervised manner because labels for degraded conditions are scarce in practice. The automation goals in monitoring can be broadly categorized as anomaly detection and recognizing the momentary fault type.

The guiding principles for algorithm selection are simplicity in terms of their straightforward visual explanation for the production managers, and the ability to progressively improve the model on the streaming data to address peculiarities in individual machine constructions.

2.4.1 Novelty detection

Anomaly, novelty, or outlier detection determines whether a health status deviates considerably from the baseline profile. The expert can then step in and diagnose the machine after the notice. Anomaly is a rare observation different from the others raising suspicion that it was created with an unrelated behavior [42]. The observations get assigned anomaly scores, and those over the threshold are novelties.

The measurements coming in the steaming fashion have to be processed in a single pass. The detection model must deal with the minimal admissible assumptions about the nature of the input events. The outliers are based on non-parametric statistical models, nearest-neighbor clustering, and isolation-based approaches [43].

DenStream is a density-based algorithm adapted from DBSCAN to cluster streaming data of arbitrarily shaped groups. Samples it includes in the first step into coherent clusters are core data points in each other's neighborhoods. Core points have at least $MinPts$ (μ) points in their neighborhood of radius Eps (ε) units. Then non-core points in the proximity area of the core point are attached to the cluster containing it [44].

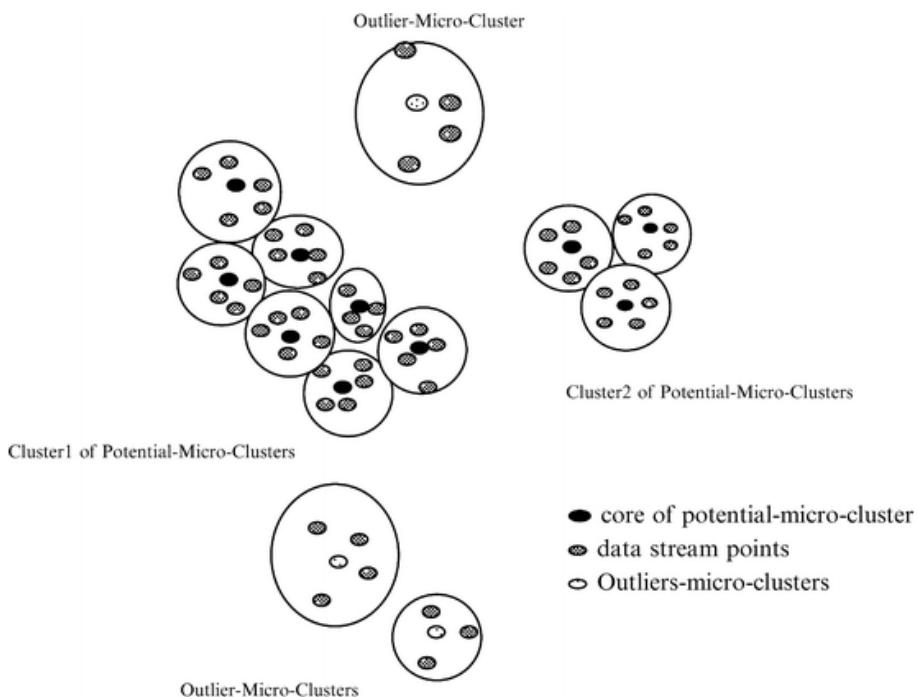


Figure 2.11: DenStream [45]

In the online maintenance phase, DenStream summarizes the nearby observations

into core *micro-clusters* that can be potential or outlier *micro-clusters* (Fig. 2.11) [46]. The (outlier) *o-micro-clusters* can grow into (potential) *p-micro-clusters* when they encompass $\beta\mu$ points. The outliers are discounted after some time in accordance to the decay function: $f(t) = 2^{-\lambda t}$ or below lower weight limit ξ . The on-demand offline stage runs DBSCAN over the approximate representation in micro-clusters to deliver final apportionment [47].

Half-Space Tree (HS-Tree) stands upon the concept of Isolation forest. It assumes that random splitting of each axis in the feature space will isolate outliers to their separate divisions sooner than non-deviant observations [43, 10]. This ensemble of trees is better suited for batch setting. HS-Tree stands out in adapting to changing streams because it is trained solely on normal data, requires constant memory, and is faster than density-based methods [48].

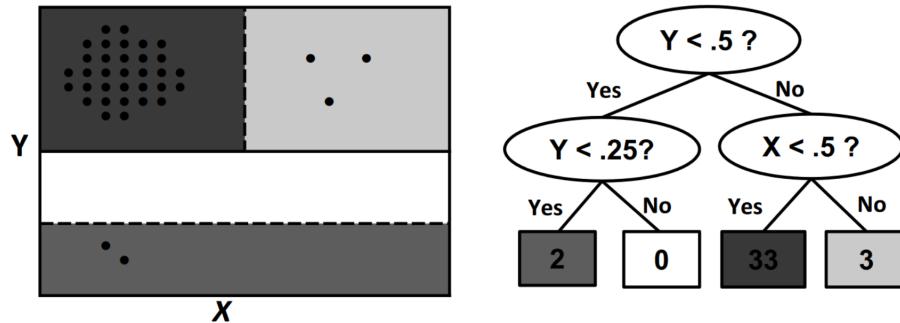


Figure 2.12: Half-space tree [48]

A full binary tree is built before the novelty detection begins by splitting tree nodes along the divisions in the randomly chosen perpendicular planes. The node stores its depth, value limits of the axis bisection (half-space), count of contained data points (mass) in two consecutive windows, and link to both child nodes [48].

The anomaly profile in the latest window is always compared to the predecessor reference window. After the latest window is filled up it replaces the reference window. It suffices to use a window size of 250 and 25 tree ensemble [48].

2.4.2 Classification

Accurate multi-class classification of machine fault causes out of the characteristics of known ones is a much more difficult task than novelty detection. Universal enough

fault baseline has to be recorded and transformed into feature space. Interactions among fault manifestations have to be accounted for. We are aware of rapid advances in knowledge transfer for deep neural networks [49]. So far, solutions seem not production ready. Therefore, we opt to use a more rudimentary model.

K-nearest neighbors (kNN) assigns the data point to the class where the majority of k closest instances belong (Fig. 2.13a). This means it can work in the semi-supervised environment because it can infer labels just from knowing a few annotations.

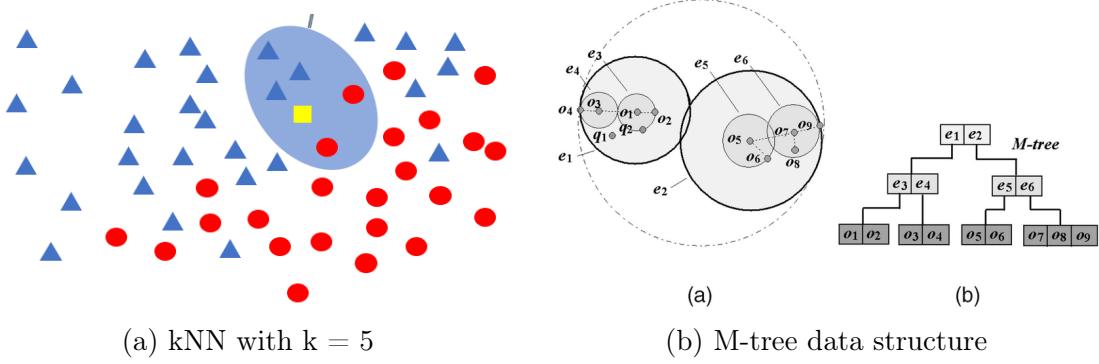


Figure 2.13: Nearest neighbors classification algorithm [50]

The sense of distance between feature vectors \mathbf{x} , \mathbf{y} have to be defined, so several metrics are available like *Euclidian distance*, *Mahalanobis distance*, or *RBF kernel* (Tab. 2.5) [51]. The demanding neighborhood queries are sped up by a Metric search tree data structure such as *M-Tree* (Fig. 2.13b). The optimal k parameter is set in supervised learning according to the breaking point in the elbow curve that plots choices of k against the error rate.

Distance	$d(\mathbf{x}, \mathbf{y})$
Euclidian distance	$\sqrt{\sum_{i=1}^n (x_i - y_i)^2}$
Mahalanobis distance	$(\mathbf{x} - \mathbf{y})^T C^{-1} (\mathbf{x} - \mathbf{y})$
Radial basis function	$\exp\left(-\frac{\ \mathbf{x}-\mathbf{y}\ ^2}{t}\right)$

Table 2.5: Distance metrics for kNN

Nearest-neighbour classifier has been successfully applied in machinery fault di-

agnostics. On the CWRU bearing dataset, the kNN with the accuracy of 96.2% slightly outperformed SVM (95%) on the combination of time and frequency-domain features, time-domain features - kNN 91.2%, SVM 88.8%, and frequency domain features - kNN (98.8%), SVM (96.2%) [52].

Comparison of kNN and KLDA on a feature set consisting of average, kurtosis, skewness, and standard deviation vectors in each domain has been conducted, achieving a data reduction rate of 95%. Best accuracies were reached for PSD features with 99.13% with KLDA and 96.64% with KNN classifiers and Mahalanobis metric. The sampling frequency was set at 40 kHz [53]. Despite kNN lagging in accuracy, we have to keep in mind annotations for faults were complete and machine learning was not tested in a streaming context.

2.5 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. [54]. The datasets listed below are also publically available online in Comma-Separated Values files.

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.

2.5.1 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. 2.14). 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 $\pm 490 \text{ m/s}^2$, and resolution step of 10.2 mV per m/s^2 .

Observations were conducted in three cardinal axes simultaneously with 2 sets

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

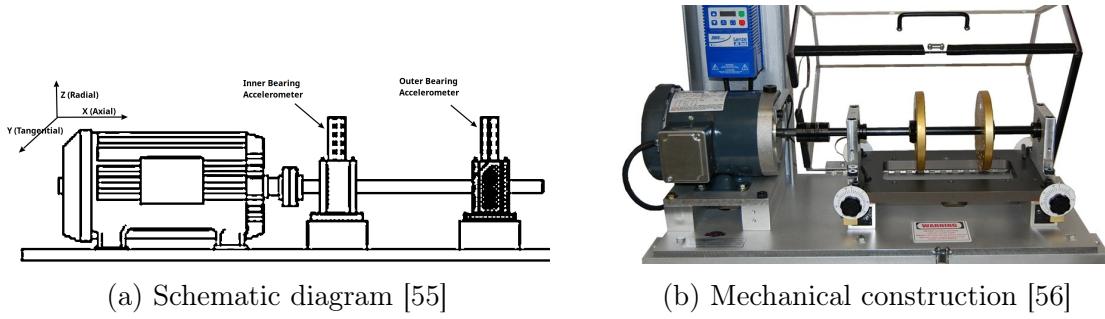


Figure 2.14: Machinery fault simulator for MaFaulDa

of accelerometers each one associated with one bearing (inner and outer bearings) (Fig. 2.14). 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.

Columns in the dataset are organized as depicted in table 2.6. 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) [55]. 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 2.6: 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 [55].

- **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 misalignment** 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 misalignment** signals were recorded under displacements: 0.50, 1.00, 1.50, 2.00 mm, each with 49 different speeds (or 50 in one instance) [55].
- **Bearing faults** are unnoticeable 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.

2.5.2 CWRU bearings dataset

In Case Western Reserve University (CWRU) bearing dataset² 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).

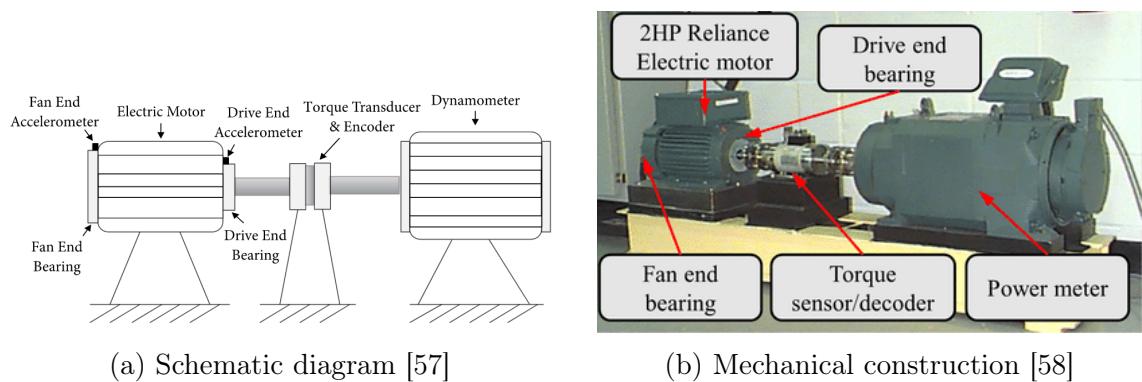


Figure 2.15: CWRU machine apparatus

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. 2.15) [52].

²<https://engineering.case.edu/bearingdatacenter/download-data-file>

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 2.7: 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. Accelerometer was sometimes mounted on the supporting base plate.

2.5.3 Unbalance of the rotating shaft

Unbalance Detection of a Rotating Shaft³ is a Kaggle dataset that simulates 4 different unbalance strengths.

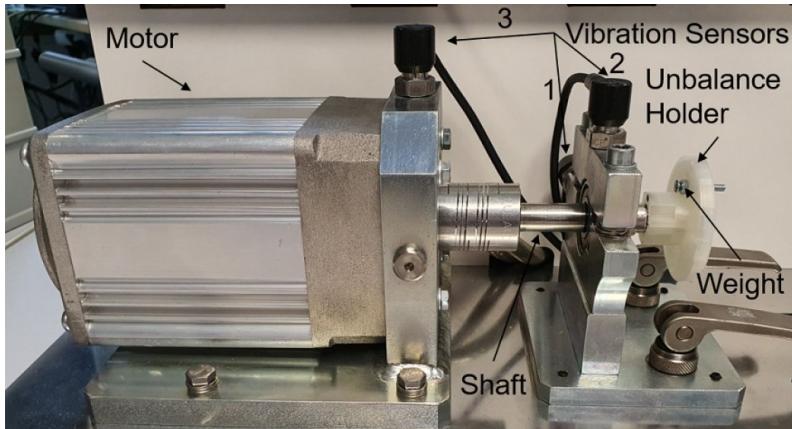


Figure 2.16: Motor driving shaft in unbalance measurement [59]

The setup is shown in Fig. 2.16. 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

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

were recorded at a sampling rate of 4 kHz [59].

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 2.8: ‘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 $\pm 490 \text{ m/s}^2$, 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 Design

In the spirit of an extensive overview of standardized procedures of industrial vibration acquisition and monitoring, the methods in feature selection, feature selection, and machine diagnostics from vibration signals, we propose a preliminary design of elements in the sensor network capable of discriminating machinery faults.

3.1 Preparatory exploration

The wood processing factory we plan to collaborate with has to choose machines that are worthy of condition monitoring. In the meantime, the data processing pipeline will be put together in Python language and packages for scientific computing. The pipeline will consist of the following stages.

1. Detrending
2. Adaptive noise cancellation of the background interference
3. Vector of all features
 - Time domain: statistical measures (Tab. 2.3)
 - Frequency domain: PSD estimation with FFT and Welch averaging with the resolution of 1 Hz combined with Hann window (Tab. 2.4)
 - Time-frequency domain: energy and entropy in wavelet coefficients from WPD and EWT filter banks.
4. Feature selection on evaluation datasets and experimental measurements from the factory to prune away irrelevant features with pearson correlation, Fisher score, and mutual information.

5. Model evaluation and comparison of novelty detection methods and precision of classification with different sets of predictors. The range of optimal parameters will be found for the DenStream (μ , ϵ , beta , λ), Half-Space Trees (window size, ensemble size), and kNN (distance metric, k neighbours). Evaluation metrics associated with confusion matrices will be used like accuracy, precision, true positive rate, and false positive rate.

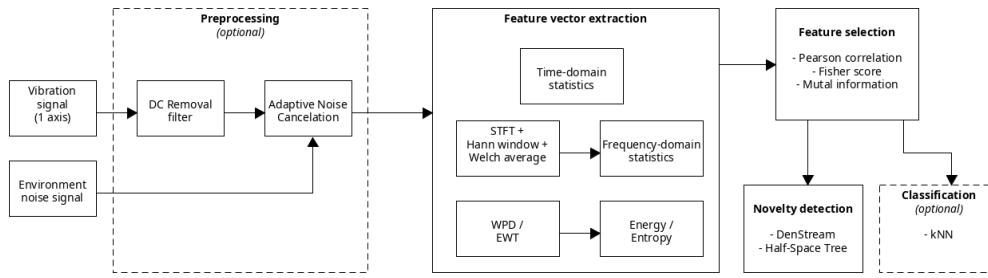


Figure 3.1: Machine learning lifecycle of feature discovery

3.2 Vibration processing

- **Input:** Samples from three-axis MEMS accelerometers, RPM tachometer, Noise background
 - **Output:** machine health status / type of fault
 - **Output on demand:** Control chart of trend features
1. **Accelerometer** - MEMS accelerometers will be placed on at least two distinct measurement points in two perpendicular axes and one sensor in the machine base for denoising purposes. Rotational speed has to be captured at the same time too. The sampling frequency shall be around 2 kHz if unbalance and looseness is to be identified, and 10 kHz if bearing faults are also of interest. The range of overall rms vibrations is not expected to exceed 30 mm/s according to the vibration severity chart.
 2. **Acquisition interval** - sensor units will be triggered in regular intervals (every 15 - 60 minutes) to collect vibration recordings from the band saw (or

another machine of choice). The machine has to be under the same load conditions every time recording is active.

3. **Features** - most relevant features are computed and compared to recent measurements. If there is a statistically significant change the whole summary is sent, otherwise keep-alive notification is sent.
4. **Wireless network** - earlier design decision has been made to establish wireless connections. Therefore, the sensor unit will send data over Wifi (IEEE 802.11), or Thread with IEEE 802.15.4 over 2.4 GHz or 868 MHz frequency bands. The application protocol shall be Constrained Application Protocol (CoAP). The messages will be encoded by Concise Binary Object Representation (CBOR) or MessagePack which provides the best compression ratio and is widely supported.
5. **Time series database** - stores history of trend values. Raw vibration measurements can be requested by the operator at any time but are available and delayed according to transfer speeds and other network constraints. The promising database technologies is TimescaleDB.
6. **Diagnosis panel** - continuously updates the anomaly detection and classification models with the introduction of annotations to notify the operator about observed faults and imminent failure of the machine. The dashboard is provided to display the current status of multiple machines.

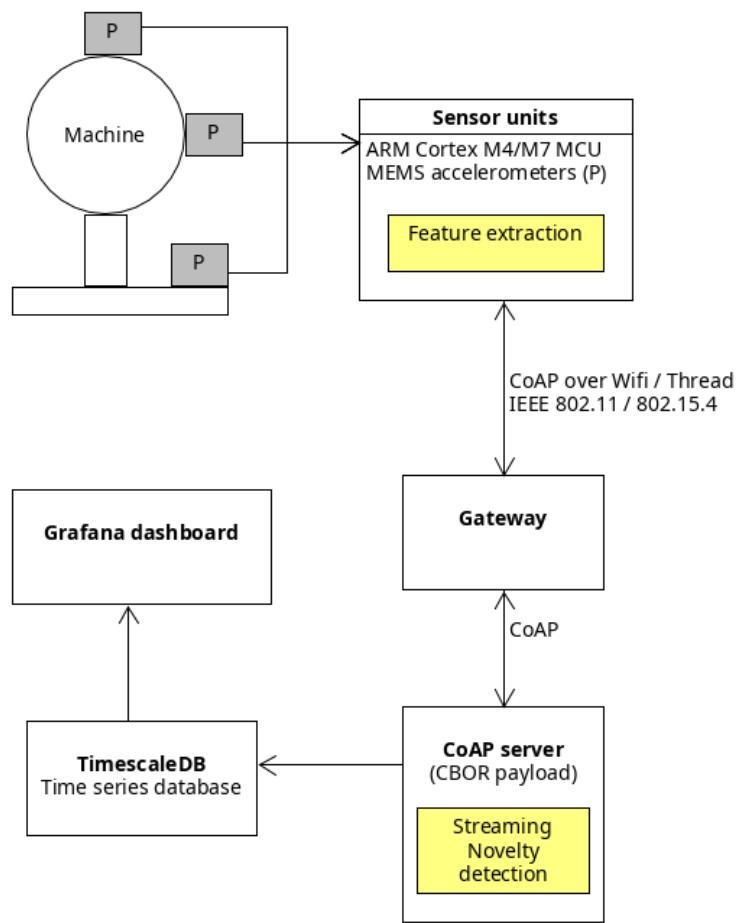


Figure 3.2: Sensor network components

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](https://doi.org/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*. 2020, vol. 12, no. 19, p. 8211. ISSN 2071-1050. Available from DOI: [10.3390/su12198211](https://doi.org/10.3390/su12198211).
5. ŽIARAN, Stanislav. *Technická diagnostika*. 1. vyd. Bratislava: Vydatelstvo 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.
7. JENNIONS, Ian K. *Integrated Vehicle Health Management: Perspectives on an Emerging Field*. SAE International, 2011. ISBN 978-0-7680-6432-2. Available from Google Books: [PIi7pwAACAAJ](https://books.google.com/books?id=7pwAACAAJ).
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](https://doi.org/10.3389/fdata.2021.644651).

LITERATURE

9. OKOH, C.; ROY, R.; MEHNEN, J.; REDDING, L. Overview of Remaining Useful Life Prediction Techniques in Through-life Engineering Services. *Procedia CIRP*. 2014, vol. 16, pp. 158–163. 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.
12. ISO 20816-1:2016 - *Mechanical Vibration - Measurement and Evaluation of Machine Vibration - Part 1: General Guidelines*. International Organization for Standardization, 2016.
13. TITTELBACH-HELMRICH, Klaus. Digital DC Blocker Filters. *Frequenz*. 2021, vol. 75, no. 9-10, pp. 331–339. ISSN 2191-6349. Available from DOI: 10.1515/freq-2020-0177.
14. LYONS, Richard G. *Understanding Digital Signal Processing*. 3rd ed. Upper Saddle River, NJ: Prentice Hall, 2011. ISBN 978-0-13-702741-5.
15. DINIZ, Paulo S. R. *Adaptive Filtering: Algorithms and Practical Implementation*. Cham: Springer International Publishing, 2020. ISBN 978-3-030-29056-6. Available from DOI: 10.1007/978-3-030-29057-3.
16. 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.
17. JOHNSON, Max Kuhn and Kjell. *Feature Engineering and Selection: A Practical Approach for Predictive Models*. 2019.

18. 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*. 2021, vol. 8, no. 4, p. 85. ISSN 2227-9709. Available from DOI: 10.3390/informatics8040085.
19. 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. ISSN 2572-6889. Available from DOI: 10.1109/ICRoM54204.2021.9663495.
20. *ISO 13373-2:2016 - Condition Monitoring and Diagnostics of Machines - Vibration Condition Monitoring - Part 2: Processing, analysis and presentation of vibration data*. International Organization for Standardization, 2016.
21. PEETERS, Geoffroy. A Large Set of Audio Features for Sound Description. 2004.
22. ADIKARAM, K.K. Lasantha Britto; HUSSEIN, Mohamed; EFFENBERGER, Matthias; 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.
23. GERBER, Timothée; MARTIN, Nadine; MAILHES, Corinne. Identification of Harmonics and Sidebands in a Finite Set of Spectral Components. 2013, vol. 1.
24. HÁJEK, Miroslav; BALÁŽ, Marcel. *Spracovanie dát generovaných senzorovou IoT sietou*. Bratislava, 2022. Bachelor's thesis. Faculty of Informatics and Information Technologies, Slovak University of Technology.
25. ARTS, Lukas P. A.; van den BROEK, Egon L. The Fast Continuous Wavelet Transformation (fCWT) for Real-Time, High-Quality, Noise-Resistant Time–Frequency Analysis. *Nature Computational Science*. 2022, vol. 2, no. 1, pp. 47–58. ISSN 2662-8457. Available from DOI: 10.1038/s43588-021-00183-z.
26. 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.

LITERATURE

27. 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*. 2022, vol. 22, no. 17, p. 6673. ISSN 1424-8220. Available from DOI: [10.3390/s22176673](https://doi.org/10.3390/s22176673).
28. 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](https://doi.org/10.1109/TIM.2019.2901514).
29. 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](https://doi.org/10.1109/TIM.2002.1017721).
30. 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](https://doi.org/10.1109/41.847906).
31. 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*. 2021, vol. 43, no. 12, p. 547. ISSN 1806-3691. Available from DOI: [10.1007/s40430-021-03259-z](https://doi.org/10.1007/s40430-021-03259-z).
32. 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*. 2022, vol. 121, no. 7, pp. 4533–4548. ISSN 1433-3015. Available from DOI: [10.1007/s00170-022-09515-3](https://doi.org/10.1007/s00170-022-09515-3).
33. GILLES, Jerome. Empirical Wavelet Transform. *IEEE Transactions on Signal Processing*. 2013, vol. 61, p. 3999. Available from DOI: [10.1109/TSP.2013.2265222](https://doi.org/10.1109/TSP.2013.2265222).
34. 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](https://doi.org/10.3390/app9091876).

35. 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](https://doi.org/10.1109/ACCESS.2020.2968851).
36. AVOCI, Moise. Spectral Negentropy and Kurtogram Performance Comparison for Bearing Fault Diagnosis. In: Dubrovnik, Croatia: International Measurement Confederation (IMEKO), 2020.
37. 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](https://doi.org/10.1007/s00170-020-05302-0).
38. 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](https://doi.org/10.1109/ACCESS.2019.2902645).
39. 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*. 2014, vol. 400, pp. 159–167. ISSN 0378-4371. Available from DOI: [10.1016/j.physa.2014.01.020](https://doi.org/10.1016/j.physa.2014.01.020).
40. 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*. 2021, vol. 502, p. 116079. ISSN 0022-460X. Available from DOI: [10.1016/j.jsv.2021.116079](https://doi.org/10.1016/j.jsv.2021.116079).
41. ZHENG, Alice; CASARI, Amanda. *Feature Engineering for Machine Learning*. O'Reilly Media, 2018. ISBN 978-1-4919-5324-2.
42. AGGARWAL, Charu C. *Outlier Analysis*. 2nd ed. Springer Publishing Company, Inc., 2016. ISBN 978-3-319-47578-3.
43. 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;

LITERATURE

- TORRE, Carmelo Maria; KARACA, Yeliz (eds.). *Computational Science and Its Applications – ICCSA 2020*. Cham: Springer International Publishing, 2020, vol. 12252, pp. 15–30. ISBN 978-3-030-58810-6. Available from DOI: 10.1007/978-3-030-58811-3_2.
44. AGGARWAL, Charu C.; REDDY, Chandan K.. *Data Clustering - Algorithms and Applications*. CRC Press, 2014. ISBN 978-1-4665-5822-9.
45. AMINI, Amineh; WAH, Teh Ying. A Comparative Study of Density-based Clustering Algorithms on Data Streams: Micro-clustering Approaches. In: *Intelligent Control and Innovative Computing*. Ed. by AO, Sio Iong; CASTILLO, Oscar; HUANG, Xu. New York, NY: Springer New York, 2012, pp. 275–287. ISBN 978-1-4614-1695-1. Available from DOI: 10.1007/978-1-4614-1695-1_21.
46. GHESMOUNE, Mohammed; LEBBAH, Mustapha; AZZAG, Hanene. State-of-the-Art on Clustering Data Streams. *Big Data Analytics*. 2016, vol. 1, no. 1, p. 13. ISSN 2058-6345. Available from DOI: 10.1186/s41044-016-0011-3.
47. 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*. Society for Industrial and Applied Mathematics, 2006, pp. 328–339. ISBN 978-0-89871-611-5. Available from DOI: 10.1137/1.9781611972764.29.
48. TAN, Swee Chuan; TING, Kai Ming; LIU, Tony Fei. Fast Anomaly Detection for Streaming Data. 2011.
49. 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.
50. CHEN, Lei; LIAN, Xiang. Efficient Processing of Metric Skyline Queries. *IEEE Trans. Knowl. Data Eng.* 2009, vol. 21, pp. 351–365. Available from DOI: 10.1109/TKDE.2008.146.
51. SHENG, Hao; CHEN, Zhongsheng; XIA, Yemei; HE, Jing. Review of Artificial Intelligence-based Bearing Vibration Monitoring. In: *2020 11th International Conference on Prog-*

- nostics and System Health Management (PHM-2020 Jinan)*. 2020, pp. 58–67. ISSN 2166-5656. Available from DOI: 10.1109/PHM-Jinan48558.2020.00018.
52. 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*. 2021, vol. 39, pp. 36–42. ISSN 2345-0533. Available from DOI: 10.21595/vp.2021.22307.
 53. 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*. 2022, vol. 22, no. 5, p. 2012. ISSN 1424-8220. Available from DOI: 10.3390/s22052012.
 54. 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*. Sociedade Brasileira de Telecomunicações, 2017. Available from DOI: 10.14209/sbrt.2017.133.
 55. PESTANA-VIANA, Denys; ZAMBRANO-LOPEZ, Rafael; de LIMA, Amaro A.; DE M. PREGO, Thiago; NETTO, Sergio L.; da SILVA, Eduardo A.B. 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)*. Florianopolis: IEEE, 2016, pp. 115–118. ISBN 978-1-4673-7835-2. Available from DOI: 10.1109/LASCAS.2016.7451023.
 56. *SpectraQuest Inc.,: Machinery Fault Simulator - Lite*. Available also from: <https://spectraquest.com/machinery-fault-simulator/details/mfs-lt/>.
 57. 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*. 2022, vol. 2022, e7214822. ISSN 1070-9622. Available from DOI: 10.1155/2022/7214822.
 58. YUHONG, Jin; HOU, Lei; CHEN, Yushu. A New Rotating Machinery Fault Diagnosis Method Based on the Time Series Transformer. 2021.
 59. MEY, Oliver; NEUDECK, Willi; SCHNEIDER, André; ENGE-ROSENBLATT, Olaf. *Machine Learning-Based Unbalance Detection of a Rotating Shaft Using Vibration Data*. 2020-07-31. Available from DOI: 10.48550/arXiv.2005.12742.

Appendix A: Plan of work

A.1 Summer semester - DP1

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

A.2 Winter semester - DP2

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