# Bachelorarbeit

## Entwicklung datengetriebener Analysemodelle zur Fehlerprävention und Ableitung von Handlungsempfehlungen beim atmosphärischen Plasmaspritzen von Zylinderkurbelgehäusen

**vorgelegt am:** 13.08.2025

**von:** Schmidt Etienne
04288 Leipzig
An der Gänseweide 3

**Studiengang:** Digital Engineering

**Studienrichtung:** -

**Seminargruppe:** DE22-1

**Matrikelnummer:** s4004953

**Praxispartner:** Leadec Automation & Engineering GmbH
09120 Chemnitz
Kauffahrtei 25 / Gebäude 1

**Gutachter:** Dipl.-Ing. (BA) Arnold, Rainer (Leadec)
Prof. Dr.-Ing. Hempel, Arne-Jens
(Staatliche Studienakademie Glauchau)

# Festsetzung des Themas der Thesis zur Bachelorarbeit

| | |
|---|---|
| Studierende(r): | Etienne Schmidt |
| Seminargruppe: | 4DE22-1 |
| Matrikelnummer: | 4004953 |
| Dualer Praxispartner: | Leadec Automation & Engineering GmbH |

**Thema der Bachelorthesis:**

Entwicklung datengetriebener Analysemodelle zur Fehlerprävention und Ableitung von Handlungsempfehlungen beim atmosphärischen Plasmaspritzen von Zylinderkurbelgehäusen

Gutachter/in Duale Hochschule: Prof. Dr.-Ing. Arne-Jens Hempel

Gutachter/in Dualer Praxispartner: Dipl.-Ing. (BA) Rainer Arnold

Herr Schmidt ist gemäß § 18 der Prüfungsordnung für den Studiengang Digital Engineering vom 01.10.2022 zur Bachelorarbeit zugelassen. Das o.g. Thema wurde am **21.05.2025** ausgegeben.

Abgabetermin der Thesis: **13.08.2025, 14:00:00 Uhr**

Prof. Ingolf Tiator
Vorsitzender
Prüfungsausschuss Technik

# Inhaltsverzeichnis

**Abbildungsverzeichnis** ..... **V**

**Tabellenverzeichnis** ..... **VII**

**Formelverzeichnis** ..... **VIII**

**Abkürzungsverzeichnis** ..... **IX**

**1 Einleitung in die Fehlererkennung mit datengetriebenen Systemen** ..... **1**

1.1 Motivation: Warum Fehlererkennung wichtig ist ..... 1

1.2 Problemstellung: Herausforderungen in der Produktionsumgebung ..... 1

1.3 Ziel: Von der Theorie zur industriellen Anwendung ..... 2

1.4 Aufbau der Arbeit: Struktur und Vorgehensweise ..... 3

**2 Prinzipieller Anwendungskontext: Atmosphärisches Plasmaspritzen in der Motorenfertigung** ..... **4**

2.1 Technologischer Hintergrund: Beschichtung durch Atmosphärisches Plasmaspritzen ..... 4

2.2 Prozessüberwachung zur Qualitätssicherung von Zylinderkurbelgehäusen ..... 6

2.3 Überwachung der Pulverinjektion ..... 7

2.4 Einflussgrößen und mögliche Fehlerquellen ..... 8

**3 Methodische Ansätze zur Erkennung von Anomalien als Grundlage der Fehlererkennung und Handlungsempfehlungsableitung** ..... **9**

3.1 Begriffliche Grundlagen und Abgrenzungen ..... 9

3.1.1 Begriff der Normalität und Anomalie ..... 9

3.1.2 Anomalie Detektion als Grundlage für die Fehlererkennung und Handlungsempfehlungsableitung ..... 10

3.1.3 Klassifizierung von Anomalien ..... 10

3.2 Methoden zur Anomalie-Erkennung ..... 13

3.2.1 Statistische Verfahren ..... 13

3.2.1.1 Standardabweichung als Streuungsmaß zur Anomaliedetektion ..... 13

3.2.1.2 Boxplots und Interquartilsabstand zur Anomaliedetektion ..... 14

3.2.2 Maschinelles Lernen ..... 15

3.2.3 Deep Learning und die Funktionsweise von Künstlichen Neuronalen Netzen ..... 18

3.3 Auswahl einer Methode zur Anomalie Erkennung ..... 22

3.4 Mathematische Beschreibung ausgewählter Verfahren ..... 23

3.4.1 Maximal Margin Classifier ..... 23

3.4.2 Support Vector Classifier ... 25
3.4.3 Support Vector Machine ... 26
3.4.4 One Class Support Vector Machine ... 27
**4 Entwicklung eines datengetriebenen Anomalie-Erkennungssystems ... 30**
4.1 Vorgehen nach dem CRISP-DM Prozess ... 30
4.2 Datensammlung und -beschreibung ... 30
4.3 Vorverarbeitung: Bereinigung, Normalisierung und Feature Engineering ... 33
4.4 Training der One Class Support Vector Machine ... 37
4.4.1 Beschriftung der Prüfdaten ... 37
4.4.2 Metriken zur Bewertung der Prognosegüte ... 38
4.4.3 Hyperparameteroptimierung ... 40
4.4.4 Evaluierung der Modellgüte ... 44
**5 Integration und Evaluation vom Anomalie-Erkennungssystem ... 49**
5.1 Planung der Modellintegration in die produktive Umgebung ... 49
5.2 Verfügbarkeit und Nutzung des Modells ... 51
5.3 Validierung und Kommentierung von erkannten Anomalien ... 53
5.4 Tests und Limitationen des entwickelten Systems ... 54
5.5 Typische Anomaliebeispiele ... 55
**6 Fazit und Ausblick ... 57**
6.1 Zusammenfassung der wichtigsten Ergebnisse ... 57
6.2 Kritische Reflexion und praxisnahe Bewertung ... 58
6.3 Weiterentwicklungsmöglichkeiten und Forschungsperspektiven ... 59
**Quellenverzeichnis ... 61**

# Abbildungsverzeichnis

**Abbildung 1** Aufbau des Brenners beim atmosphärischen Plasmaspritzen ..... 5
**Abbildung 2** Aufbau eines Regelkreises ..... 7
**Abbildung 3** Uni- und Multivariate Anomalien ..... 12
**Abbildung 4** Punkt- und Gruppenanomalien mit und ohne zeitlichen Kontext..... 12
**Abbildung 5** Normalverteilung mit Wahrscheinlichkeitsintervallen..... 14
**Abbildung 6** Boxplot ..... 15
**Abbildung 7** Regression und Klassifikation mit Anomalien ..... 17
**Abbildung 8** Künstliches Neuron mit ReLU-Aktivierungsfunktion ..... 19
**Abbildung 9** Aktivierungsfunktionen für ein Neuron..... 20
**Abbildung 10** Aufbau eines neuronalen Netzes ..... 20
**Abbildung 11** Maximal Margin Classifier ..... 24
**Abbildung 12** Slackvariable $\xi$ ..... 25
**Abbildung 13** Maximal Margin Classifier und Support Vector Classifer bei Daten mit Ausreißern ..... 25
26
**Abbildung 14** Lineare Trennung im höherdimensionalen Raum ..... 26
**Abbildung 15** CRISP-DM-Prozess ..... 30
**Abbildung 16** Zeitreihenausschnitt der Pulverlinienmerkmale..... 31
**Abbildung 17** Kontextuelle Anomalie im Kontext der Bypassschaltung..... 32
**Abbildung 18** Datenvorverarbeitungsprozess ..... 33
**Abbildung 19** Transformierte Daten und Einteilung in Teilsequenzen..... 34
**Abbildung 20** Methode des gleitenden Fensters ..... 35
**Abbildung 21** Minimum-Maximum Streudiagramm der Stichproben von Merkmal 1 ..... 38
**Abbildung 22** Raum der Gittersuche ..... 41
**Abbildung 23** Kreuzvalidierung mit k = 10 Teilmengen ..... 42
**Abbildung 24** Hyperparameteroptimierung von „nu“ und „gamma“ mittels Gittersuche..... 43
**Abbildung 25** Histogramm der Entscheidungsfunktionswerte ..... 44
**Abbildung 26** Abstände der falsch negativen Beobachtungen zur Hyperebene.... 45
**Abbildung 27** Abstände der falsch positiven Beobachtungen zur Hyperebene..... 45
**Abbildung 28** Abstände der richtig positiven Beobachtungen zur Hyperebene..... 45
**Abbildung 29** Anpassung des Grenzwertes der OCSVM zur Performanceoptimierung ..... 46
**Abbildung 30** Falsch negativ klassifizierte Daten (grün) und falsch beschriftete Daten (rot)..... 47
**Abbildung 31** Falsch negative Beobachtungen in Bezug auf Merkmal 3 ..... 47

**Abbildung 32** Boxplot der Abstände zur Trennebene ..... 48
**Abbildung 33** Prozess der Anomaliedetektion..... 50
**Abbildung 34** Markierte Anomalie im Dashboard ..... 52
**Abbildung 35** Kommentier- und Quittierung von klassifizierten Anomalien ..... 53
**Abbildung 36** Vom AD-System erkannte Anomalien ..... 56

# Tabellenverzeichnis

**Tabelle 1** Abbildungen des Eingabevektors in höherdimensionalen Raum .... 27
**Tabelle 2** Attribute und Datentypen des Datensatzes ..... 33
**Tabelle 3** Verzögerungsspalte der Bypassstellung ..... 34
**Tabelle 4** Filterung der Datenpunkte mit inaktivem Bypass ..... 34
**Tabelle 5** Konfusionsmatrix im Kontext der Anomaliedetektion ..... 39
**Tabelle 6** Ergebnisse der Gittersuche mit Kreuzvalidierung ..... 42
**Tabelle 7** Konfusionsmatrix der Trainingsergebnisse ..... 45
**Tabelle 8** Konfusionsmatrix nach angepasstem Grenzwert ..... 46
**Tabelle 9** Konfusionsmatrix nach angepassten Labeln ..... 48
**Tabelle 10** Vergleich Polling und Interrupt ..... 49

# Formelverzeichnis


**Formel 1** Anomalien im Sinne d er Wahrscheinlichkeitstheorie ..... 10
**Formel 2** z-Transformation ..... 13
**Formel 3** Interquartilsabstand eines Boxplots ..... 15
**Formel 4** Menge der Anomalien mit Boxplots ..... 15
**Formel 5** Künstliches Neuron mit Aktivierungsfunktion ..... 19
**Formel 6** Fehlervektor ..... 21
**Formel 7** Kostenfunktion in Form des mittleren quadratischen Fehlers ..... 21
**Formel 8** Anpassung des Parametervektor ..... 21
**Formel 9** Anpassung der Netzparameter ..... 21
**Formel 10** Optimierungsproblem des Maximal Margin Classifier ..... 24
**Formel 11** Umgestelltes Optimierungsproblem des Maximal Margin Classifer ..... 24
**Formel 12** Optimierungsproblem Support Vector Classifer mit Soft Margin ..... 25
**Formel 13** Linear Kernel ohne Hyperparameter ..... 27
**Formel 14** Polynomial Kernel mit Hyperparametern $r, d$ ..... 27
**Formel 15** Radial Base Function Kernel mit Hyperparameter $\gamma$ ..... 27
**Formel 16** Primal Problem der OCSVM ..... 27
**Formel 17** Lagrange Funktion der OCSVM ..... 28
**Formel 18** Duales Problem der OCSVM ..... 28
**Formel 19** Entscheidungsfunktion der OCSVM ..... 28
**Formel 20** Reduktion der Größe des Datensatzes durch Merkmalsextraktion ..... 36
**Formel 21** Fehler ..... 39
**Formel 22** Genauigkeit ..... 39
**Formel 23** Präzision ..... 39
**Formel 24** Sensitivität ..... 39
**Formel 25** Fβ -Maß ..... 39

# Abkürzungsverzeichnis

| | |
|---|---|
| **AD** | Anomalie Detektion |
| **CRISP-DM** | Cross Industrie Standard Process for Data Mining |
| **DL** | Deep Learning |
| **IQR** | Interquartilsabstand |
| **KNN** | Künstliches Neuronales Netz |
| **ML** | Maschinelles Lernen |
| **OCSVM** | One Class Support Vector Machine |
| **ReLU** | Rectified Linear Unit |
| **SVM** | Support Vector Machine |
| **tanh** | Hyperbolischer Tangens |
| **ZKG** | Zylinderkurbelgehäuse |

# 1 Einleitung in die Fehlererkennung mit datengetriebenen Systemen

## 1.1 Motivation: Warum Fehlererkennung wichtig ist

Die zunehmende Digitalisierung und Vernetzung industrieller Anlagen im Zuge der Industrie 4.0 verändert die Art und Weise wie technische Produktionsprozesse überwacht und optimiert werden. Eine zentrale Rolle spielt hierbei die Prozessüberwachung. Sie hilft Unregelmäßigkeiten im Prozessablauf frühzeitig zu erkennen.

Die frühzeitige Fehlererkennung ist in modernen Industrieprozessen wichtig. Sie trägt zur Reduktion von Ausschuss, Minimierung von Stillstandzeiten und der Sicherstellung der Produktqualität bei. Datengetriebene Modelle sind neben physikalisch- und wissensbasierten Modellen die meistverbreitete Variante neuester Entwicklungen im Bereich der vorausschauenden Instandhaltung.[1]

Insbesondere die Anomaliedetektion (AD) hat sich als ein effektives Mittel zur Identifikation unerwarteter oder seltener Prozesszustände etabliert. Für die AD ist kein explizites Vorwissen über mögliche Fehlerbilder nötig. Sie ermöglicht die frühzeitige Warnung vor potenziellen Störungen und ist somit ein grundlegendes Werkzeug der datengetriebenen Überwachung.[2]

## 1.2 Problemstellung: Herausforderungen in der Produktionsumgebung

Trotz der möglichen Potentiale datengetriebener Modelle zur AD stellt die reale Produktionsumgebung verschiedene Anforderungen und Herausforderungen, welche bei der Integration und Nutzbarmachung beachtet werden müssen. Die Integration datengetriebener AD in bestehendem laufendem Prozess, wie etwa der Beschichtungsprozess von Zylinderkurbelgehäusen (ZKG), wird deshalb durch mehrere Faktoren erschwert.

Ein Problem ergibt sich aus der Heterogenität der Prozesse. Verschiedene Maschinen, Produkte und Betriebsmittelzustände führen zu variierenden Datenmustern. Verschiedene Umwelteinflüsse verändern diese zusätzlich. Dies erschwert die

[1] Vgl. ZONTA; u.a., 2020 S. 1 f.
[2] Vgl. CHANDOLA; BANERJEE; KUMAR, 2009 S.2

Generalisierbarkeit von Modellen, weshalb adaptive oder trainierbare Ansätze erforderlich sind.[3]

Zentrales Problem ist auch die Datenverfügbarkeit. Es kann sein, dass es bei selten auftretenden Fällen keine oder nur wenige historische Beispiele gibt. Zudem fehlen in vielen Fällen präzise beschriftete Daten, welche den Einsatz von überwachten Modellen (vgl. Kapitel 3.2.2) und im Allgemeinen die Bewertung der Modellgüte erschweren.

Hinzu kommt, dass viele Produktionsprozesse zeitvariant sind, dass heißt sie ändern sich im Zeitverlauf durch Wartung, Materialchargen, Umweltbedingungen oder wechselnde Konfigurationen. Dies birgt die Gefahr, dass ein Modell seine Gültigkeit verliert und deshalb angepasst, weiter optimiert und neu trainiert werden muss.

Schließlich ist auch die Interpretierbarkeit der Ergebnisse von großer Bedeutung. Neben dem Ergebnis selbst sind auch seine Darstellung und Nutzbarmachung entscheidend. Das System soll dem Werker eine Möglichkeit bieten, Rückmeldung zu gefundenen Anomalien zu geben, um mögliche Handlungsempfehlungen ableiten zu können und erfasste Anomalien weiter zu klassifizieren.

Daraus ergibt sich folgende zentrale Problemstellung für diese Arbeit:

> Wie kann ein robustes und rückmeldefähiges System zur AD in der Produktionsumgebung des Prozesses zur Beschichtung von ZKG‘s implementiert werden? Es soll nicht nur zuverlässig Abweichungen zum Normalverhalten erkennen, sondern auch am laufenden Prozess durch Werker validiert und so weiter optimiert werden können. Aus den gefundenen Prozessanomalien sollen mögliche Handlungsempfehlungen abgeleitet werden können.

## 1.3 Ziel: Von der Theorie zur industriellen Anwendung

Ziel dieser Arbeit ist es, einen praxisnahen Ansatz zur AD zu entwickeln und in eine reale Produktionsumgebung zu integrieren. Das System soll robust mit Sensordaten aus einem Beschichtungsprozess umgehen können und eine Schnittstelle zur Rückmeldung von Fachpersonal enthalten.

Der Fokus dieser Arbeit liegt dabei die Auswahl und Optimierung eines Analyseverfahrens aus dem Bereich des Maschinellen Lernens. Neben der Modellierung und der

[3] Vgl. MUNIR; u.a., 2019 S. 9

Datenverarbeitung selbst stehen auch die Visualisierung, Interpretierbarkeit und Nutzerintegration im Vordergrund dieser Arbeit.

Mit diesem System soll die Grundlage für weitere Analysen geschaffen werden. Ziel ist es, basierend auf gefundene Anomalien, diese zu gruppieren und zu beschreiben. Damit soll die Grundlage für eine automatische Handlungsempfehlungsableitung bei typischen Verhalten geschaffen werden.

Diese Arbeit verfolgt damit einen interdisziplinären Ansatz mit dem Ziel, ein anwenderortientiertes AD-System bereitzustellen, dass diesen Anforderungen gerecht wird.

## 1.4 Aufbau der Arbeit: Struktur und Vorgehensweise

Die vorliegende Arbeit ist in 6 Kapitel gegliedert. Ziel ist es schrittweise von dem Produktionsprozess, über die theoretischen Grundlagen datengetriebener Anomalie-erkennung bis zur praktischen Umsetzung eines Prototypens in einer realen Produktionsumgebung vorzudringen. Der Aufbau folgt der Struktur des Cross Industrie Standard Process for Data Mining. (CRISP-DM)

*Kapitel 2* beschreibt den konkreten Anwendungsfall, das atmosphärische Plasmaspritzen in der Motorenfertigung. Neben einem technischen Überblick über das Produktionssystem, werden die Anforderungen an die Prozessüberwachung, relevante Datenquellen und potenzielle Fehlerquellen erläutert.

*Kapitel 3* liefert die theoretischen Grundlagen der AD. Weiterhin werden sowohl die Bereiche der Statistik, des maschinellen Lernens (ML) und des Deep Learning (DL) zur AD vorgestellt. Anschließend erfolgt eine mathematische Beschreibung ausgewählter Modelle, mit den Fokus auf Support Vektor Verfahren, mit dem Ziel die One Class Support Vector Machine (OCSVM), als ein ML-Verfahren, darzulegen.

*Kapitel 4* behandelt die praktische Entwicklung eines AD-Systems. Das Vorgehen orientiert sich an den Phasen 2-5 des CRISP-DM-Prozessmodells und umfasst die Datenerhebung, -vorverarbeitung, Modellentwicklung, Hyperparameteroptimierung und Evaluierung des Modells.

*Kapitel 5* dokumentiert die Integration und Evaluation des entwickelten Systems in der realen Produktionsumgebung. Neben der technischen Implementation werden Möglichkeiten zur Kommentierung erkannter Fehler bzw. Anomalien entwickelt. Abschließend erfolgt die Bewertung der Systemleistung im laufenden Prozess.

*Kapitel 6* fasst zum Schluss, alle wesentlichen Ergebnisse zusammen, Bewertet die erarbeiteten Ansätze kritisch und gibt einen Ausblick auf mögliche Weiterentwicklungen und offene Forschungsfragen.

# 2 Prinzipieller Anwendungskontext: Atmosphärisches Plasmaspritzen in der Motorenfertigung

Dieses Kapitel beschreibt den industriellen Anwendungskontext dieser Arbeit. Der Fokus liegt hierbei auf dem atmosphärischen Plasmaspritzen (APS), als Beschichtungstechnologie im Kontext der Motorenfertigung, insbesondere bei der Beschichtung von ZKG. APS wird eingesetzt um Schichten mit speziellen Eigenschaften, wie zum Beispiel Verschleißfestigkeit und Temperaturbeständigkeit, aufzubringen.

## 2.1 Technologischer Hintergrund: Beschichtung durch Atmosphärisches Plasmaspritzen

Es gibt eine Vielzahl an thermischen Spritzverfahren mit unterschiedlichen Eigenschaften.[4] Das APS ist ein solches thermisches Spritzverfahren, bei dem ein Prozessgas oder Gasgemisch durch einen elektrischen Lichtbogen auf relativ hohe Temperaturen gebracht und ionisiert wird.[5]

Zwischen zwei Elektroden wird hierfür eine hohe Gleichspannung angelegt. Bei ausreichender Energiezufuhr werden Elektronen aus den Atomen des Prozessgases gelöst. Durch die Ionisation entsteht schließlich ein Plasmastrahl mit Temperaturen bis zu 20.000°C. Schräge Bohrungen am Auslass des Brenners sorgen hierbei für einen Drall des Gases und eine Einschnürung des Lichtbogens. Bei diesem Vorgang wird elektrische Energie in thermische und kinetische Energie umgewandelt.[6]

Der entstehende Plasmastrahl schmilzt das Beschichtungsmaterial in Pulverform auf und beschleunigt es mit relativ hoher Geschwindigkeit Richtung Werkstückoberflächen.

Die Entstehung des Plasmas ist abhängig von[7]:

- Elektrodenabstand
- Elektrodenform
- Potentialdifferenz / Spannung

Der Beschichtungsprozess benötigt verschiedene Eingabematerialien und -medien, darunter zählen:

[4] Vgl. online: Oerlikon Group, 2024 (04.08.2025)
[5] Vgl. DIN EN ISO 14917
[6] Vgl. RICHTER, 2014 S. 15 f.
[7] Vgl. PASCHEN, 1889 S. 36

- Energie in Form von Elektrizität (zur Erzeugung des Plasmas)
- Prozessgase (z.B. Argon oder Wasserstoff) als Plasma und Transportmedium
- Spritzwerkstoff (Pulvermaterial)
- Kühlflüssigkeit (z.B. Wasser oder Luft) zur Prozessstabilisierung

Der Spritzwerkstoff wird im Plasmastrahl aufgeschmolzen und durch die Strömungsgeschwindigkeit des Strahls in Richtung Bauteil beschleunigt. Dabei zerfällt das Pulver in feine kleine Tropfen, trifft die Bauteiloberfläche und bildet nach relativ schnellen erstarren eine Schicht. Um den gesamten Zylinder gleichmäßig zu beschichten, bewegt sich der Brenner in den Zylindern des Bauteils mittels Robotersteuerung vertikal und rotierend. In Abbildung 1 ist der Aufbau eines Brenners des APS schematisch dargestellt.

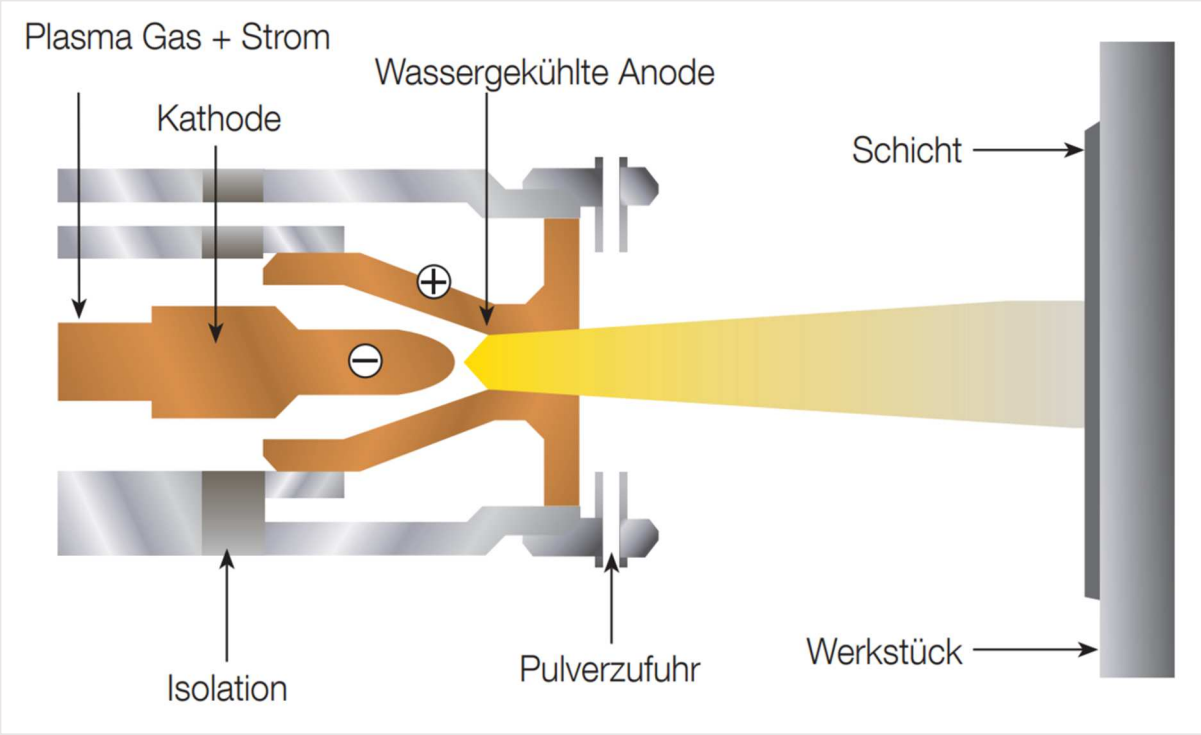



**Abbildung 1** Aufbau des Brenners beim atmosphärischen Plasmaspritzen[8]

[8] online: Oerlikon Group, 2024 (04.08.2025)

## 2.2 Prozessüberwachung zur Qualitätssicherung von Zylinderkurbelgehäusen

Für die beschichteten Bauteile muss sichergestellt werden, dass diese innerhalb vorgegebener Qualitätsgrenzen liegen. Um dies zu gewährleisten, ist eine ausreichende Prozessstabilität erforderlich. Prozessanomalien sind dabei häufig ein Hinweis auf instabile und fehlerhafte Abläufe. Fehler im Prozess frühzeitig zu erkennen, trägt dazu bei, Abweichungen der Bauteile von den Qualitätsvorgaben zu reduzieren.

Zudem können viele Qualitätsmerkmale eines Bauteils oft nur stichprobenartig bestimmt werden, da ihre Bestimmung im vollen Umfang eine Zerstörung des gesamten Bauteils nach sich zieht. Deshalb ist es wichtig neben den Qualitätsgrößen auch die relevanten Prozessgrößen zu überwachen, da diese einen Einfluss auf die Qualität haben.

Einerseits können die Prozessgrößen mit Absolutwerten überwacht werden. Dabei wird geprüft, ob ein Wert innerhalb definierter Toleranzgrenzen liegt. Die Grenzwertüberwachung ist insbesondere bei stationär verhaltenden Prozessgrößen sinnvoll, bei denen sich die Größe stabil um einen Idealwert bewegt.[9].

Viele der Größen im Beschichtungsprozesses sind geregelt. Das bedeutet sie unterliegen einem Regelkreis, der auf den Prozess einwirkende Störungen kompensiert. Er ist essenziell, um den Prozess stabil zu halten. Er erschwert jedoch die Erkennung von Störsignalen, da sie in der Prozessgröße nicht immer sichtbar werden.[10]

Beispielsweise könnten Betriebsmittelkomponenten leicht beschädigt oder beeinträchtigt sein. Aufgrund der Kompensierung durch den Regelkreis kann es sein, dass diese Beeinträchtigung jedoch nicht in den Prozessgrößen direkt auffällt.

In Abbildung 2 ist ein solcher Regelkreis dargestellt. Die Führungsgröße $w$ stellt eine Sollwertvorgabe für die Regelgröße $y$ dar. Eine Abweichung $e$ zwischen $w$ und der gemessenen Prozessgröße $y_m$ führt zur Bestimmung der Stellgröße $u$. Die Stellgröße wirkt auf die Prozessstrecke ein und beeinflusst dadurch $y$. Besonders hervorzuheben ist hierbei die Fehlerabweichung e, da sie einen Vergleich zwischen Soll- und Istwert darstellt und das System rückkoppelt. Störungen, welche intern & extern auftreten können dadurch weitestgehend kompensiert werden.

[9] ISERMANN, 2006 S. 377 f.
[10] ISERMANN, 2006 S. 377 f.

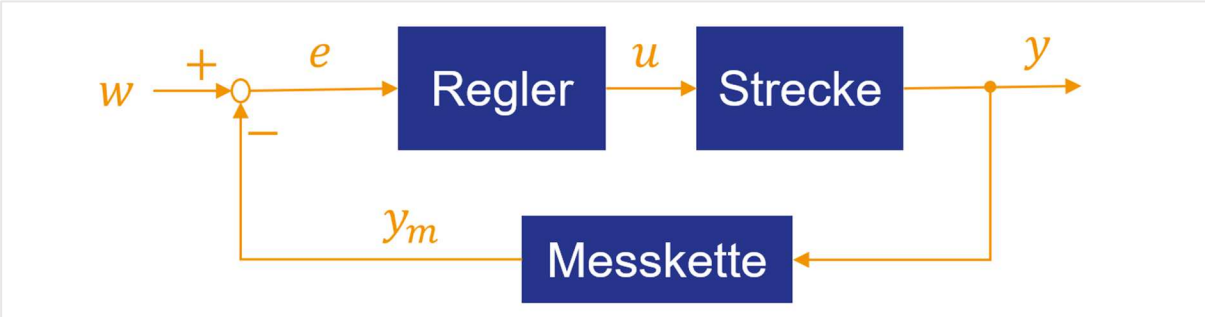



**Abbildung 2** Aufbau eines Regelkreises [11]

Weiterhin können Anomalien innerhalb des Toleranzbereiches auftreten oder sind erst durch die Betrachtung mehrerer Größen zueinander sichtbar. Daher ist eine reine Absolutwertüberwachung oft nicht ausreichend, um potenzielle Probleme und Veränderungen zu erkennen. Es bedarf einer genaueren Detektion, die auch kontext-behaftete, multivariate und versteckte Störungen erkennt.

Um solche Störungen frühzeitig zu erkennen, ist es notwendig, weitere Größen in die Überwachung mit einzubeziehen. Hier können zusätzliche Signale, wie zum Beispiel die zur Regelgröße gehörigen Stellgröße helfen Störungen und Fehler zu identifizieren. Um Störungen im Beschichtungsprozess über die Absolutwertüberwachung hinaus zuverlässig zu erkennen ist ein Verständnis über die relevanten Systeme und Prozessgrößen erforderlich

## 2.3 Überwachung der Pulverinjektion

Der APS-Prozess lässt sich in mehrere funktionalen Systeme gliedern[12]:

- Pulver und Pulverinjektion
- Gasversorgung
- Robotersteuerung zur Brennerbewegung
- Energieversorgung
- Kühlung

Im Rahmen dieser Arbeit liegt der Fokus auf der Überwachung der Pulverinjektion, d.h. der Förderung des Beschichtungsmaterials vom Pulvertopf bis zur Injektion des Pulvers in den Plasmastrahl des Brenners. Die Regelung der Fördermenge erfolgt über einen Fördermotor. Die Pulverförderung stellt einen zentralen Aspekt für die Prozessqualität dar, da Abweichungen in der Fördermenge, Einspritzgeschwindigkeiten und Betriebsmittelzuständen Einfluss auf die Qualität der Beschichtung haben.

[11] Eigene Abbildung (vgl. HEMPEL MSR Kapitel 1.4)
[12] Vgl. RICHTER, 2014 S. 27 f.

Die Überwachung dieses Teilsystems soll potenzielle Fehler und Abweichungen frühzeitig identifizieren. Zwar bieten auch andere Teilsysteme (Energieversorgung, Kühlung, …) Potenzial für eine Überwachung über eine Absolutwertüberwachung hinaus. Aufgrund des Umfangs dieser Arbeit liegt der Fokus jedoch nur auf der Pulverinjektion als relevantestes System neben der Energieversorgung und Stromregelung.

## 2.4 Einflussgrößen und mögliche Fehlerquellen

Die Qualität der Pulverinjektion hängt stark vom Zustand der Betriebsmittelkomponenten und Umweltbedingungen ab. Faktoren wie Düsenverschleiß, Leckagen und Oberflächenbeschaffenheiten der Förderschläuche beeinflussen unter anderem die:

- Fördermenge des Pulvers
- Partikelflugbahn
- Aufschmelzrate der Partikel
- Verteilungen im Partikelstrom

Ein deutlicher Einfluss hat die Wechselwirkung zwischen Fördermenge und Plasmatemperatur. Eine zu hohe Pulvermenge kann die Plasmatemperatur abkühlen und so die Aufschmelzrate der Partikel reduzieren. Gleichzeitig erhöht sich durch größere Partikelkonzentrationen die Wahrscheinlichkeit von Partikelkollisionen, wodurch sich Tröpfchen zusammenschließen und so ihre Form und Masse beeinflusst wird.[13]

Typische Fehlerquellen in der Pulverförderung sind unter anderem:

- Verstopfte Düse
- Defekter oder undichter Förderschlauch
- Leeres Pulversystem
- Mechanische Einwirkungen auf das System

Solche Muster lassen sich in den Prozessdaten wiederfinden. Die Daten von Sensoren, welche Prozessgrößen wie Fördermengen, Drücke und Fördermotorengeschwindigkeiten messen, enthalten relativ selten solche Fehlerbilder.

Die zuvor dargestellten Rahmenbedingungen und potenziellen Fehlerquellen verdeutlichen die Notwendigkeit einer zuverlässigen Überwachung der Pulverinjektion. Im folgenden Kapitel werden deshalb die theoretischen Grundlagen und methodischen Ansätze der datengetriebenen AD, als Überwachungswerkzeug der Pulverinjektion, untersucht.

[13] Vgl. RICHTER, 2014 S. 19, 25

# 3 Methodische Ansätze zur Erkennung von Anomalien als Grundlage der Fehlererkennung und Handlungsempfehlungsableitung

Die Fähigkeit unerwartete oder abweichende Prozesszustände zu erkennen, bildet die Voraussetzung für die Fehlererkennung. Die AD bildet hierbei die methodische Grundlage, um abweichende Muster zu erkennen, unabhängig davon, ob es sich um einen auf die Qualität des Produktes auswirkenden Fehlers handelt.

Ziel dieses Kapitels ist es, die methodischen Grundlagen und Varianten zur AD zu erläutern. Zunächst werden dafür die Begriffe der Normalität und Anomalie erläutert. Anschließend wird ein Überblick über verschiedene Varianten der datengetriebenen AD gegeben. Von statistischen über klassischen Verfahren des ML bis hin zum DL. Ein besonderer Schwerpunkt bildet hierbei die OCSVM, welche im praktischen Teil als zentrales Modell untersucht wird.

## 3.1 Begriffliche Grundlagen und Abgrenzungen

### 3.1.1 Begriff der Normalität und Anomalie

Die Definition von „Normalität“ ist stark kontextabhängig und anwendungsspezifisch. In technischen Prozessen lässt sich die Normalität nicht unabhängig vom betrachteten System und gegebener Nebenbedingungen definieren. Was als Normal gilt, ergibt sich aus dem stabilen Verhalten eines Prozesses in einem gewissen Zeitraum. [14]

Man kann nicht davon ausgehen, dass Normalität an einer Maschine identisch mit der einer anderen Maschine ist, auch wenn dasselbe System zugrunde liegt. Zum Beispiel können unterschiedliche Längen von Schläuchen zu anderen gemessenen Werten der Prozessgrößen führen. Werte, welche an einer Maschine normal sind, können an einer weiteren Maschine von der Normalität abweichendes Verhalten darstellen.

**Definition:** „Eine *Anomalie* ist ein Fall oder eine Gruppe von Fällen, die in irgendeiner Weise ungewöhnlich sind und nicht zu den allgemeinen Mustern passen, die von der Mehrheit der Daten gezeigt werden.“ [15]

Anomalien sind also einzelne Datenpunkte oder Gruppen an Datenpunkten, die selten sind und ungewöhnliches Verhalten des Prozesses widerspiegeln. Sie stellen eine

[14] Vgl. WINTER S. 236
[15] FOORTHUIS, 2021 S. 300; Übersetzt

Vielfalt von der Normalität abweichendem Verhalten dar. Anomalien können statisch o. dynamisch auftreten.[16]

Anomalien sind im Sinne der Wahrscheinlichkeitstheorie alle Beobachtungen, welche in einem Bereich geringerer Wahrscheinlichkeit zum restlichen Datensatz fallen. Wenn $X \subseteq \Re^D$ der Eingabevektorraum eines technischen Prozesses ist, in dem die Beobachtungen $x_i$ in einer Wahrscheinlichkeitsverteilung $P^+$ von $X$ liegen, dann ist die Menge der Anomalien als Teilmenge von $X$ über einen Grenzwert $\tau$ wie folgt definiert[17]:

$$A = \{x \in X \,|p^+(x) \leq \tau\} \mid \tau \geq 0 \,, X \subseteq \Re^D$$

**Formel 1** Anomalien im Sinne d er Wahrscheinlichkeitstheorie[18]

### 3.1.2 Anomalie Detektion als Grundlage für die Fehlererkennung und Handlungsempfehlungsableitung

Ein ungewöhnlicher Prozesszustand heißt jedoch nicht unmittelbar, dass dieser auch einen Fehler im Prozess darstellt. Es kann anormales Verhalten auftreten, welches sich nicht negativ auf Qualität des Produktes oder den Betriebsmittelzustand auswirkt. Die Erkennung von Anomalien gehört hierbei zur deskriptiven Datenanalyse. Es wird lediglich festgestellt, ob die Daten normal oder ungewöhnlich sind.

Dagegen sind Fehlererkennungssysteme und solche zur Handlungsempfehlungsableitung eher der diagnostischen Analyse zuzuordnen, da sie Ursachen und Auswirkungen mitberücksichtigen müssen. Ein System zur AD bildet jedoch die Grundlage für die weitere diagnostische Analyse.

### 3.1.3 Klassifizierung von Anomalien

In der Literatur wird oft zwischen Anomalien, Ausreißern und Neuheiten (engl.: „novelities“) differenziert. Die Begriffe werden oftmals auch Synonym benutzt. Ausreißer sind meist einzelne Datenpunkte, die sich statistisch stark vom normalen Unterscheiden.[19]

Einzelne Datenpunkte, die von der Norm abweichen, werden auch als Punktanomalien bezeichnet. Im Gegensatz zu Ausreißern umfassen Anomalien auch komplexere Abweichungen, welche kontextbezogen und in Gruppen auftreten können. Ein

[16] Vgl. FOORTHUIS, 2021 S. 297 f.
[17] Vgl. RUFF, 2021 S. 759
[18] RUFF, 2021 S. 759
[19] Vgl. HODGE; AUSTIN, 2004 S. 85

Ausreißer ist also eine Anomalie, wogegen eine Anomalie jedoch kein Ausreißer sein muss.

Neuheiten sind unbekannte bzw. neue Datenmuster und können bei zeitvarianten Systemverhalten auftreten. Sie können einerseits fehlerhaftes Verhalten und andererseits neue Bereiche von Normalität darstellen.[20]

Alle drei Formen befinden sich in Bereichen geringer Wahrscheinlichkeit der Wahrscheinlichkeitsverteilung $P^+$ des Eingabevektorraumes $X$. In dieser Arbeit werden alle dieser Arten betrachtet und im Folgenden wird der Begriff der Anomalie synonym benutzt. [21]

Anomalien werden typischerweise in 3 Arten unterteilt:

- Punkt- und Gruppenanomalien
- Kontextbezogene Anomalien (auch bedingte Anomalien)

Im technischen Prozess werden meist viele verschiedene Größen gemessen. Eine Prozessgröße kann hierbei eine andere beeinflussen. Tritt eine Anomalie in einer Größe auf, ist diese univariat. Dagegen sind multivariate Anomalien erst in Betrachtung mehrerer Größen zusammen erkennbar. (vgl. Abb. 3)

Eine multivariate Zeitreihe lässt sich in einer Matrize X darstellen. Sie enthält die gemessen Werte von K Sensoren k zu N diskreten Zeitpunkten n:

**Definition[22]:**

$$X = \big(x(1), x(2), \dots x(N)\big)^T = \begin{pmatrix} x_1(1) & \cdots & x_1(N) \\ \vdots & \ddots & \vdots \\ x_K(1) & \cdots & x_K(N) \end{pmatrix}$$

$$x(n) = \big(x_1(n), x_2(n), \dots, x_3(n)\big)$$

Für eine univariate Zeitreihe entspricht $x(n)$ einem vom Sensor zum Zeitpunkt n gemessen skalaren Wert. Eine univariate Zeitreihe lässt sich also als ein Vektor darstellen.

Von einer Punktanomalie spricht man, wenn sich eine einzelne Beobachtung vom Rest der Daten unterscheiden. Sind mehrere Datenpunkte zusammen ungewöhnlich, so spricht man von Gruppenanomalien. Tritt eine Anomalie erst im Kontext auf (z.B. in Bezug auf Zeit oder Ort), dann ist diese bedingt. (vgl. Abb. 4) [23]

[20] Vgl. RUFF, 2021 S. 760
[21] Vgl. RUFF, 2021 S. 760
[22] BELAY; u.a., 2023 S.3
[23] Vgl. CHANDOLA; BANERJEE; KUMAR, 2009 S.7 ff.

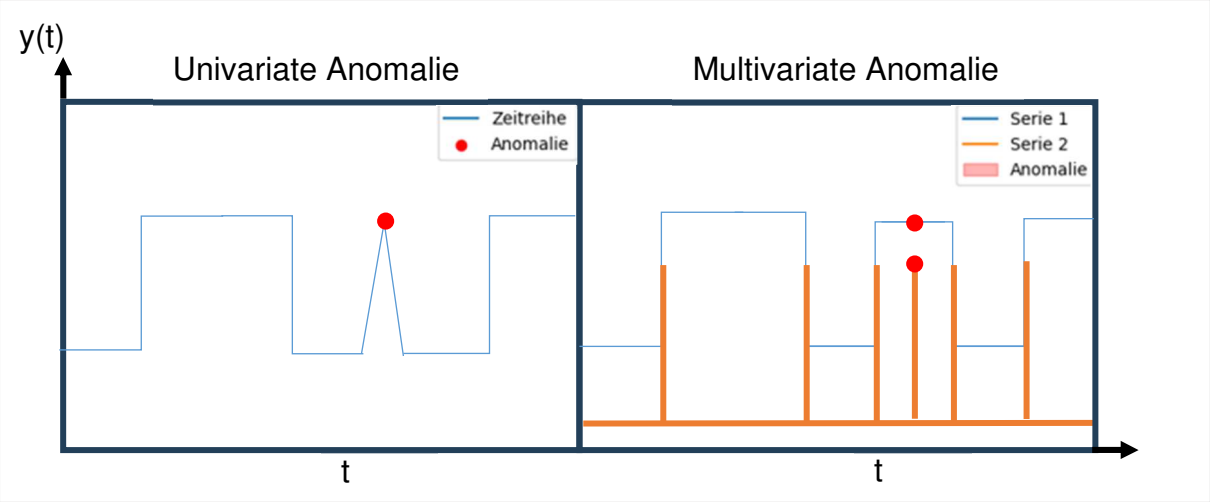



**Abbildung 3** Uni- und Multivariate Anomalien[24]

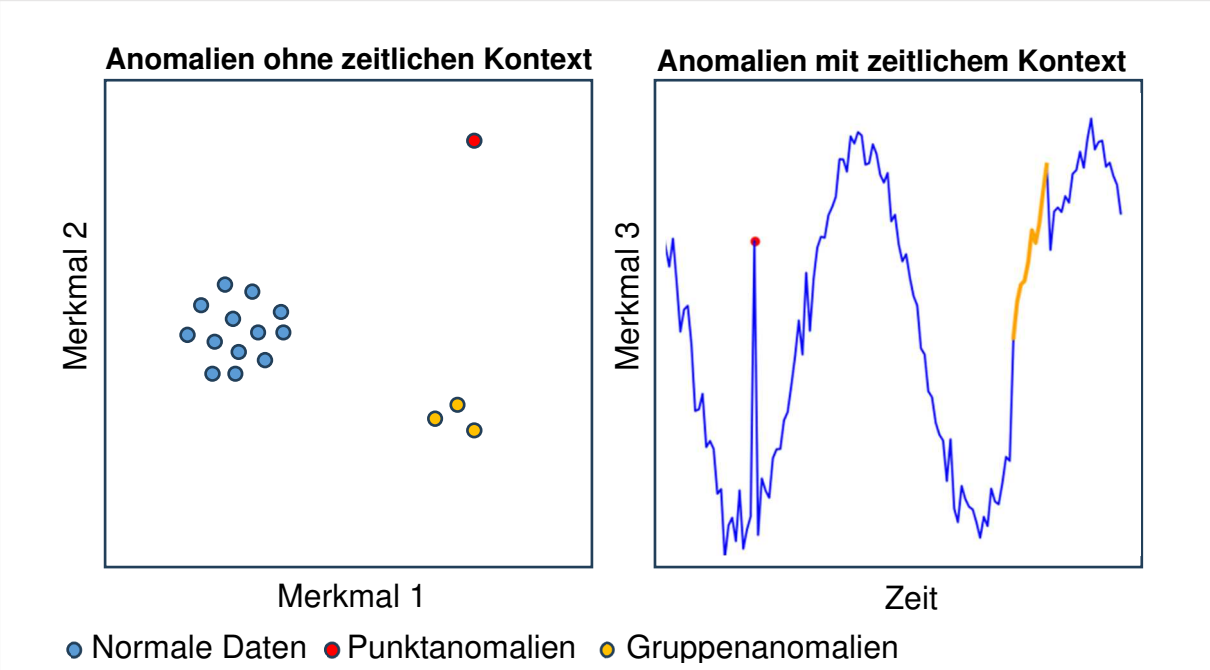



**Abbildung 4** Punkt- und Gruppenanomalien mit und ohne zeitlichen Kontext[25]

Im folgenden Kapitel werden verschiedene Bereiche des „Data Mining“ vorgestellt, um die verschiedenen Arten von Anomalien zu erkennen.

[24] Eigene Abbildung
[25] Eigene Abbildung

## 3.2 Methoden zur Anomalie-Erkennung

### 3.2.1 Statistische Verfahren

Statistische Methoden sind erste Ansätze mit denen Ausreißer in Datensätzen gefunden werden können. [26] Wie in *Kapitel 3.1.1* geschildert, liegen anormale Instanzen in Bereichen niedriger Wahrscheinlichkeit des Datensatzes. Ziel ist es diese Beobachtungen zu identifizieren.

Es können verschiedene statistische Maße wie zum Beispiel die Standardabweichung, Varianz, Spannweite, Median, Mittelwert und Kovarianzen ermittelt werden.

Zu den Mitteln der statistischen univariaten AD gehören zum Beispiel:

- Standardabweichung von Standardnormalverteilungen
- Boxplots

In den nächsten Kapiteln werden beide Methoden näher erläutert.

#### 3.2.1.1 Standardabweichung als Streuungsmaß zur Anomaliedetektion

Standardnormalverteilungen sind Normalverteilungen, welche einen Mittelwert von 0 und einer Standardabweichung von 1 besitzen. Jede Normalverteilung, die dies nicht mit sich bringt, kann durch eine z-Transformation standardisiert werden. (vgl. Formel 2) [27]

$$z_i = \frac{x_i - \bar{x}}{\sigma_x}$$

**Formel 2** z-Transformation[28]

$\sigma$…Standardabweichung, $\bar{x}$…Mittelwert der Normalverteilung

Werte der Standardnormalverteilung welche größer $\tau$ oder kleiner als $-\tau$ werden hierbei, als Ausreißer klassifiziert. Wie groß der Grenzwert $\tau$ gewählt wird hängt vom Kontext ab. Je größer $\tau$, desto weniger Datenpunkte werden als Anomalien klassifiziert. Je nach Grenzwert verändert sich die Wahrscheinlichkeit des Intervalls.[29]

[26] Vgl. HODGE; AUSTIN, 2004 S. 86
[27] Vgl. WITTE, 2019 S. 20, 33
[28] Vgl. WITTE, 2019 S. 33
[29] online: LOHNINGER, 2021 (11.07.2025)

Für verschiedene standardisierte Normalverteilungen sind die Wahrscheinlichkeiten für dasselbe Intervall gleich. (vgl. Abbildung 5) Sind die Daten nicht Normalverteilt gelten diese Wahrscheinlichkeiten nicht.

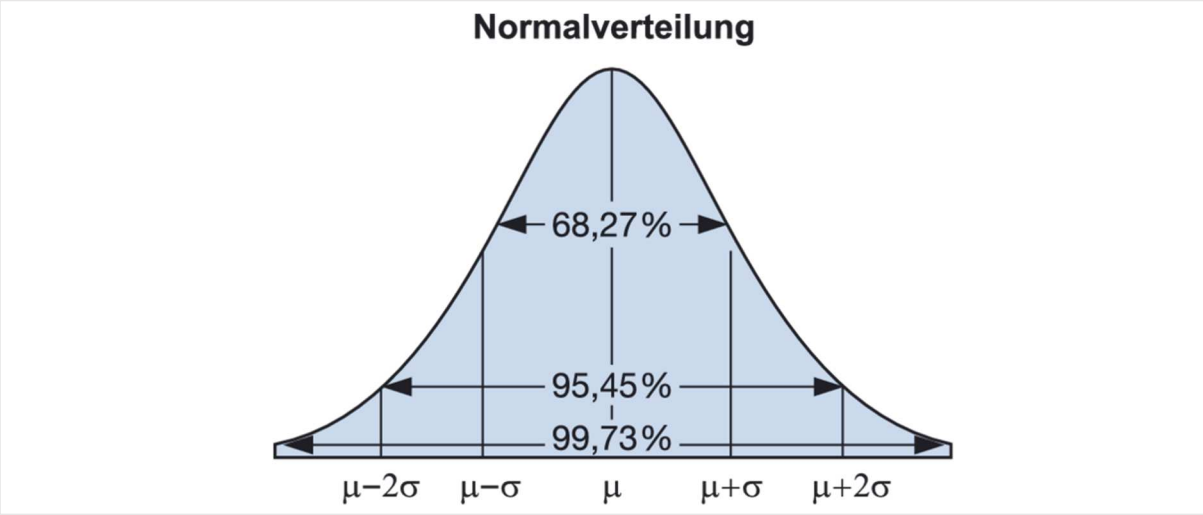



**Abbildung 5** Normalverteilung mit Wahrscheinlichkeitsintervallen[30]

### 3.2.1.2 Boxplots und Interquartilsabstand zur Anomaliedetektion

Eine weitere Methode zur Identifikation univariater Ausreißer sind Boxplots. Boxplots sind ein graphisches Werkzeug, um die Verteilungen eines numerischen Merkmals zu visualisieren. (vgl. Abbildung 6)

Zunächst ordnet man die Liste mit dem gemessenen Werten der Größe nach. Danach werden die Quantile wie folgt ermittelt:

- Das erste Quantil ($Q_1$) entspricht dem Wert an der 25%-Stelle
- Das zweite Quantil ($Q_2$), also der Median, entspricht dem Wert in der Mitte (50%-Stelle)
- Das dritte Quantil ($Q_3$) entspricht dem Wert an der 75%-Stelle

Aus der Differenz zwischen $Q_3$und $Q_1$ergibt sich der Interquartilsabstand (IQR). (vgl. Formel 3) Aus dem IQR und den Quartilen $Q_1$und $Q_3$. ergeben sich die Whisker, welche die Ober und Untergrenze für die Ausreisererkennung festlegen. Typischerweise haben die Whisker eine Länge des 1,5-Fachen IQR. [31]

Die Menge der Anomalien o. Ausreißer nach der Boxplot Methode ergibt sich aus Formel 4.

[30] KAMPS, 2018
[31] Vgl. LAURIKKALA; JUHOLA; KENTALA, 2000 S.21

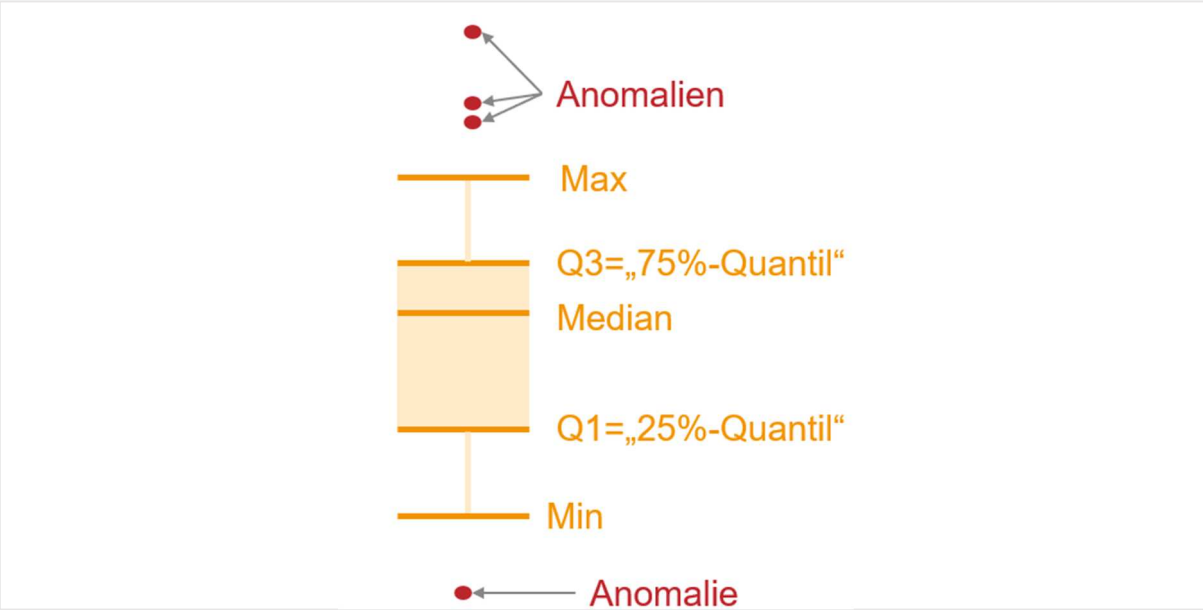



**Abbildung 6** Boxplot[32]

$$IQR = Q_3 - Q_1$$

**Formel 3** Interquartilsabstand eines Boxplots

$$A = \{x \in X \mid Q_1 - 1{,}5 * IQR > \ a > Q_3 + 1{,}5 * IQR\}$$

**Formel 4** Menge der Anomalien mit Boxplots[33]

Die Statistik liefert die Grundlagen für die Analyse der Daten. Sie ist ein Werkzeug, um Zusammenhänge zu verstehen und kann eingesetzt werden, um einfache Anomalien im Datensatz zu erkennen. Jedoch stößt sie bei komplexeren Mustern und Zusammenhänge und varianten Prozessverhalten an ihre Grenzen.

Die erläuternden Methoden bilden die Grundlage für das maschinelle Lernen (ML). Im folgenden Kapitel wird das Gebiet des ML und dessen Methoden zur AD näher betrachtet.

### 3.2.2 Maschinelles Lernen

ML erweitert die Prinzipien der Statistik um algorithmische Modelle. Diese sind in der Lage automatisch Modelle auf Basis eines Algorithmus zu erstellen, um Prognosen und Prädiktionen durchzuführen, meist ohne, dass explizit Annahmen getroffen werden müssen.[34]

[32]Eigene Abbildung vgl. CHANDOLA; BANERJEE; KUMAR, 2009 S. 34
[33]Vgl. LAURIKKALA; JUHOLA; KENTALA, 2000 S. 21
[34]Vgl. AKSHAY B.; u.a., 2025 S.14 ff.

Grundlage ist ein abstrahiertes Modell mit freien Parametern. Ziel ist es diese freien Parameter anzupassen, sodass ein konkreter Zusammenhang zwischen Ein- und Ausgabe eines Datensatzes hergestellt wird. Der Anpassungsprozess der Parameter wird als Modelltraining bezeichnet.[35]

Einerseits können Werte einer Zielgröße prädiziert werden. Die mathematischen Modelle sind Abbildungen, die die Eingabevektoren $x \in \mathfrak{R}^n$ auf $y \in \mathfrak{R}$ abbilden. Die Daten können auch diskretisiert vorliegen. Hierbei spricht man von einen Regressionsproblem. Regression kann zur AD genutzt werden, indem der Fehler zwischen der Ausgabe des Modells und eines realen Prozessmesswertes überwacht wird. Ist ein Fehler zu groß kann der Datenpunkt als Anomalie klassifiziert werden. Regressionsmodelle sind jedoch selbst oft gegen solche Ausreißer anfällig. Die Ausreißer sollten vor dem Trainingsprozess aus dem Datensatz entfernt werden.[36]

Andererseits bestimmt die Klassifikation diskrete Klassen. Ziel hierbei ist es eine Trennlinie zwischen 2 oder mehreren Klassen zu finden. Die Eingabe wird auf eine diskrete Ausgabe $y \in C$ abgebildet, wobei $C = \{c_1, c_2, \dots, c_n\}$ die Menge aller n definierten Klassen ist. Im Kontext der AD entspricht $C = \{\mathrm{Normal}, \mathrm{Anormal}\}$. Hierbei kann in Ein-Klassen-Modelle und Multi-Klassen-Modelle unterschieden werden. Multi-Klassen-Modelle basieren auf der Annahme, dass der Datensatz mit mehreren Labeln versehen ist. Dagegen werden Ein-Klassen-Modelle nur mit Daten eines Labels, z.B. den normalen Instanzen trainiert.[37]

Abbildung 7 zeigt sowohl die Regression und Klassifikation anhand einfacher Beispiele. Während bei der Regression ein Zusammenhang zwischen der Eingangsgrößen und den Ausgangsgrößen gesucht wird, wird bei der Klassifikation versucht die Klassen mittels der Eingabedaten voneinander zu trennen. In der Abbildung ist schematisch eine lineare Klassifikation (links) und eine lineare Regression (rechts) dargestellt. Dabei können Ausreißer (rot) das Modell beeinflussen.

Weiterhin können in einem Datensatz, welcher keine Zielwerte enthält, Ähnlichkeiten und Gruppen mit Clusteringverfahren identifiziert werden.

Zu den Methoden zur AD aus dem Bereich des ML zählen zum Beispiel:

- Isolation Forest
- Support Vector Machine
- K-Means
- Bayes Netz

[35]Vgl. ALPAYDIN, 2022 S. 1
[36]Vgl. CHANDOLA; BANERJEE; KUMAR, 2009 S. 35 f.
[37]Vgl. CHANDOLA; BANERJEE; KUMAR, 2009 S.20 f.

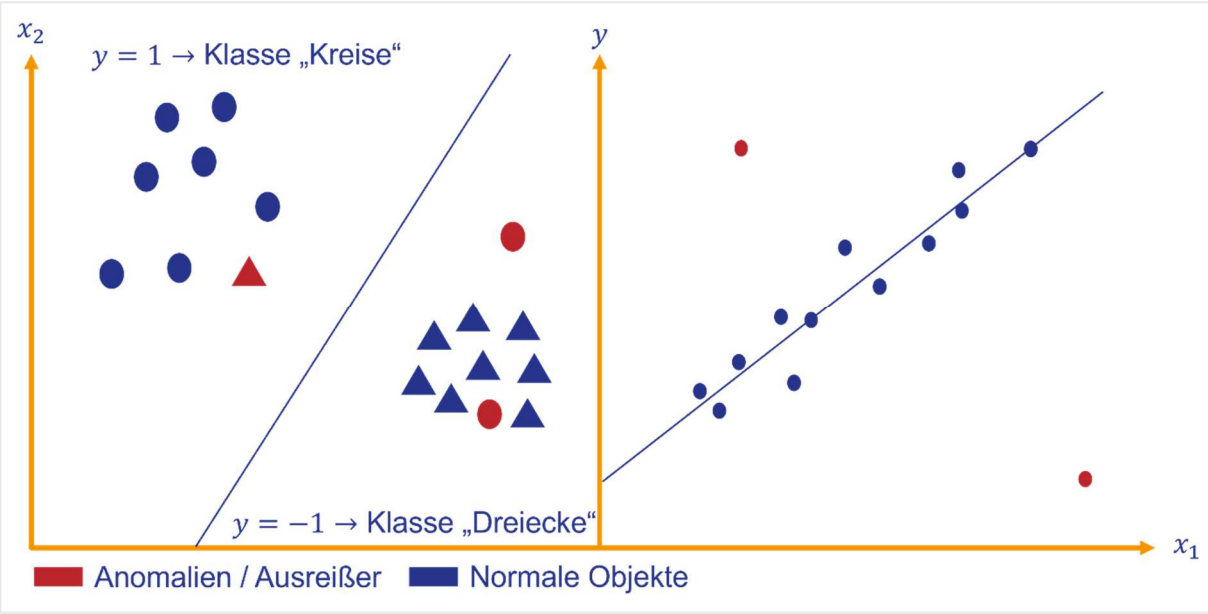



**Abbildung 7** Regression und Klassifikation mit Anomalien [38]

Weiterhin lassen sich die Methoden des ML in 4 Lernarten unterteilen, welche von der Art des verwendeten Modells, der Verfügbarkeit und Struktur der Trainingsdaten abhängig sind:

1) Überwachtes Lernen

Die Modelle brauchen einen Datensatz der sowohl die Eingabedaten (Merkmale) als auch die dazugehörigen Zielwerte (Label) enthält. Ziel ist es eine Abbildung der Merkmale auf die Label zu erlernen, damit unbekannte Daten klassifiziert werden können.[39]

2) Unüberwachtes Lernen

Im unüberwachten Lernen werden ausschließlich nicht gelabelte Daten verwendet. Ziel ist die Struktur-, Muster- oder Gruppenerkennung. Diese Methoden werden oft zur AD eingesetzt, da selten gelabelte Fehlerdaten vorliegen. [40]

3) Selbstbestärkendes Lernen

Hierbei lernen die Modelle durch Interaktion mit ihrer Umwelt. Es erhält Rückmeldungen, in Form von Belohnungen oder Bestrafungen, basierend auf der Güte der getroffenen Entscheidungen und passt sich so während des Prozesses an.[41]

[38] Eigene Abbildung vgl. ALPAYDIN, 2022 S. 6, 11
[39] Vgl. GOODFELLOW; COURVILLE; BENGIO, 2016 S. 103
[40] Vgl. GOODFELLOW; COURVILLE; BENGIO, 2016 S. 103
[41] Vgl. GOODFELLOW; COURVILLE; BENGIO, 2016 S. 103

4) Semiüberwachtes Lernen

Hierbei werden gelabelte, sowie nicht gelabelte Daten für das Modelltraining verwendet. Hierbei kann zum Beispiel nur eine Klasse für einen Teil des Datensatzes bekannt sein. Halbüberwachtes Lernen wird oft da eingesetzt, wo die Beschriftung der Daten zeitaufwendig ist, aber eine große Menge nicht beschrifteter Daten vorliegen.[42]

### 3.2.3 Deep Learning und die Funktionsweise von Künstlichen Neuronalen Netzen

DL erweitert den Ansatz des maschinellen Lernens, indem es künstlichen neuronale Netze (KNN) behandelt. Diese KNN haben viele versteckte Schichten (deshalb „Deep“) zwischen der Eingabe- und Ausgabeschicht.

KNN sind an den organischen Prozessen von Nervenzellen inspirierte mathematische Modelle. Sie bestehen aus Schichten miteinander verbundener Neuronen.

Verschiedene Typen von KNN lassen sich durch folgende 3 Merkmale unterscheiden[43]:

5) Neuronen
6) Verbindungen
7) Lernregeln

Diese drei Merkmale bestimmen die Struktur und Funktionsweise des KNN. Wie in Abbildung 8 dargestellt, besteht das klassische künstliche Neuron meist aus mehreren Eingängen $x_i$ welche mit den Gewichten $w_{i,j}$ multipliziert werden. Die gewichteten Eingänge werden zusammen mit dem Bias $w_{0,j}$ aufsummiert. Die Ausgabe des Neurons wird durch die Aktivierungsfunktion bestimmt.

Wie in Formel 5 zu sehen, lässt sich das künstliche Neuron mathematisch mit einer Summennotation darstellen. Um diese zu vereinfachen, fasst man die Eingänge $x_{1,j} \dots x_{n,j}$ und Wichtungen $w_{1,j} \dots w_{n,j}$ des Neurons $j$ als Vektoren $x_j, w_j \in \Re^n$ zusammen und ersetzt den Bias $w_{0,j}$ mit $b_j$. Der Ausdruck lässt sich nun mit dem Skalarprodukt aus $x_j$ und $w_j$ darstellen.[44]

$$x_j = f\left(w_{0,j} + \sum_{i=1}^{n} w_{i,j} * x_{i,j}\right) = f\left(w_j \cdot x_j + b_j\right)$$

[42] Vgl. online: Was ist halbüberwachtes Lernen? | IBM, 2025 (05.08.2025)
[43] Vgl, HEINE; u.a., 2023 S. 27 f.
[44] Vgl. HEMPEL TBPI Kapitel 4.1.1

**Formel 5** Künstliches Neuron mit Aktivierungsfunktion [45]

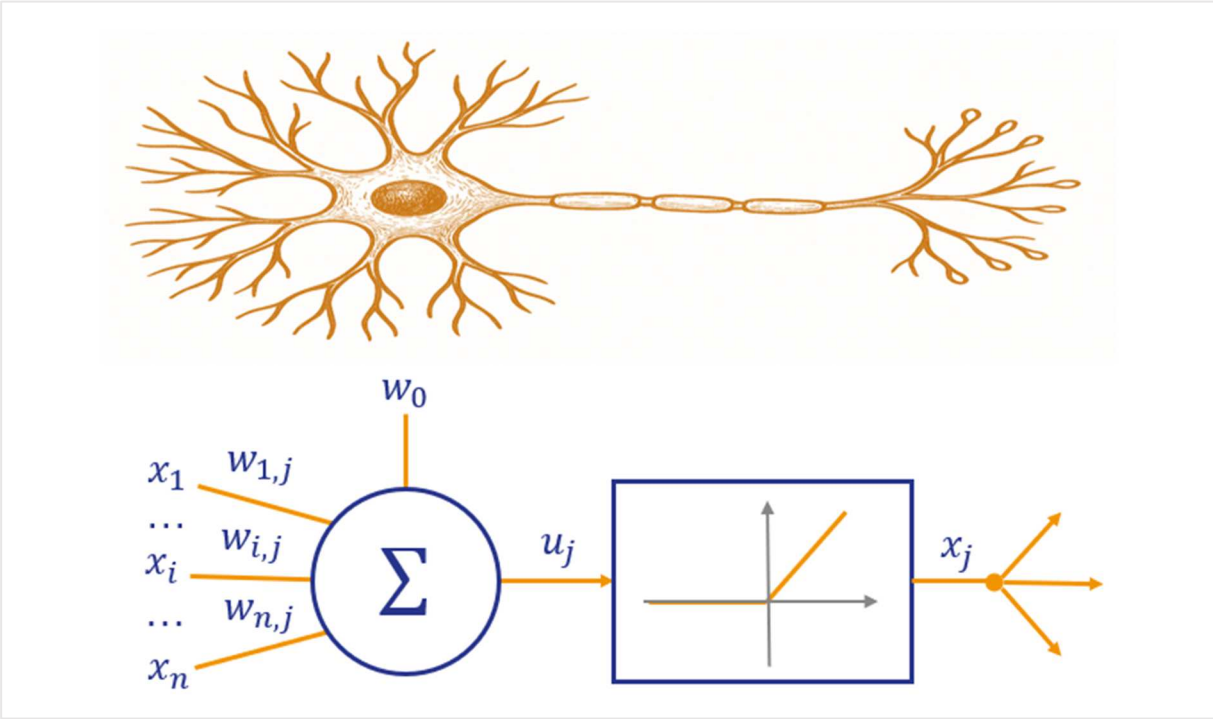



**Abbildung 8** Künstliches Neuron mit ReLU-Aktivierungsfunktion [46]

Beispielhafte Aktivierungsfunktionen (vgl. Abbildung 9) sind:

- Heaviside Funktion
- Rectified Linear Unit (ReLU)
- Sigmoid Funktion
- Hyperbolischer Tangens (tanh)

Mehrere Neuronen werden zu einem Netz zusammengefasst und sind Schicht für Schicht miteinander verbunden. Abbildung 10 zeigt ein einfaches KNN mit einer versteckten Schicht. Die Neuronen dieser Schicht sind mit jedem Neuron der Eingangs- und Ausgangsschicht verbunden. In diesem Fall spricht man von einer vollständig vernetzen Schicht. [47]

[45] Vgl. NIELSEN, 2015 S.4,6,13
[46] Eigene Abbildung (vgl. Formel 5)
[47] Vgl. online: Hochschule Trier S.8

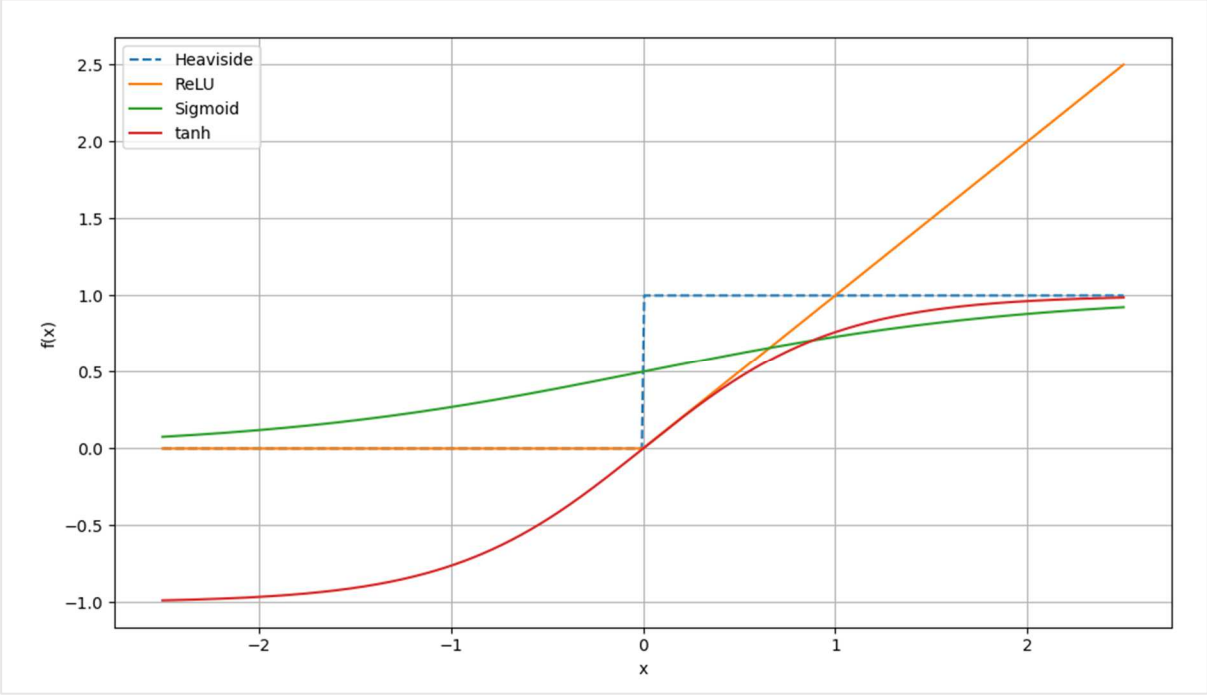



**Abbildung 9** Aktivierungsfunktionen für ein Neuron[48]

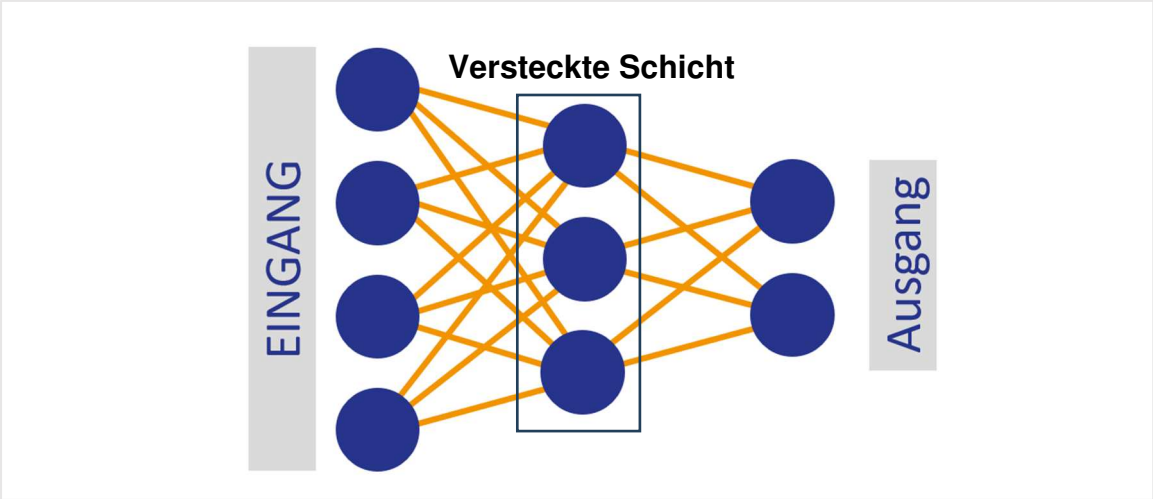



**Abbildung 10** Aufbau eines neuronalen Netzes[49]

Das Ziel des Netzes ist es eine Abbildung zwischen den Ein- und Ausgabedaten auf Grundlage des Trainingsdatensatz zu erlernen. Dabei können ausschließlich die Gewichte und der Bias der Neuronen angepasst werden, um eine Beziehung möglichst optimal zu modellieren. Man kann sie als Parametervektor $\theta$ zusammenfassen. [50]

Zu Beginn des Trainingsprozesses wird das KNN zunächst mit zufälligen Werten initialisiert. Anschließend schätzt das Netz für einen Eingabevektor $u$ ein Ausgabevektor

[48] Eigene Abbildung
[49] Eigene Abbildung
[50] Vgl HEMPEL TBPI Kapitel 4

$y_m$. Der Fehlervektor $e$ ergibt sich aus der Differenz zwischen der Modellausgabe $y_m$ und dem gewünschten Zielausgabevektor $y_p$. (vgl. Formel 6)

$$e = y_m - y_p$$

**Formel 6** Fehlervektor[51]

Die Kostenfunktion C – häufig in Form des mittleren quadratischen Fehlers (vgl. Formel 7) wird verwendet, um den Abstand der gewünschten Ausgabe von der des Netzes zu messen. Ziel ist es diese Funktion zu minimieren.

$$C(\theta) = \frac{1}{2}\sum_{i=1}^{n}\left(y_{m,i}(\theta) - y_{p,i}\right)^2 = \frac{1}{2}\|e(\theta)\|^2$$

**Formel 7** Kostenfunktion in Form des mittleren quadratischen Fehlers [52]

Zur Anpassung der Gewichte wird das Netz nach bestimmten Lernalgorithmen trainiert. Ein gängiges Verfahren ist die Kombination der Backpropagation in Verbindung mit dem Gradientenabstiegverfahren. Ziel ist es die Parameter so zu optimieren, dass eine möglichst genaue Beziehung zwischen Ein- und Ausgabe angelernt wird.

Hierfür wird der Gradient der Kostenfunktion bestimmt, welcher die Richtung und Steigung des größten Anstieges angibt. Der negative Gradient dagegen gibt die Richtung des größten Abstiegs an. Der neue Parametervektor ermittelt sich aus der Subtraktion des alten Parametervektor und dem Produkt aus dem negativen Gradienten der Kostenfunktion und der Lernrate.[53] (vgl. Formel 8)

$$\theta^{(t+1)} = \theta^{(t)} - \alpha * \nabla \mathrm{C}(\theta)$$

**Formel 8** Anpassung des Parametervektor[54]

Für jeden neuen Parameter bedeutet dies, die partielle Ableitung der Kostenfunktion zu dem Parameter zu bestimmen, mit der Lernrate zu multiplizieren und von dem alten Parameterwert zu subtrahieren. (vgl. Formel 9)

$$w_{ij}^{(t+1)} = w_{ij}^{(t)} - \alpha * \frac{\partial}{\partial w_{ij}} C(\theta)$$

$$b_j^{(t+1)} = b_j^{(t)} - \alpha * \frac{\partial}{\partial b_j} C(\theta)$$

**Formel 9** Anpassung der Netzparameter[55]

[51] Vgl. HEMPEL TBPI Kapitel 4
[52] Vgl. HEMPEL TBPI Kapitel 4
[53] Vgl. KASTE, 2024 S. 69 ff.
[54] Vgl. KASTE, 2024 S. 69 ff., 76
[55] Vgl. KASTE, 2024 S. 69 ff., 76

Künstliche Neuronale Netze können hierbei auf verschiedene Weisen auf Basis von Regression oder Rekonstruktion zur AD eingesetzt werden.

In den folgenden Kapiteln wird ein datengetriebenes Verfahren für die AD der Pulverinjektion des Beschichtungsprozesses ausgewählt und näher beschrieben.

## 3.3 Auswahl einer Methode zur Anomalie Erkennung

Für die AD stehen eine Vielzahl an datengetriebenen Verfahren zur Verfügung. Wie in den vorherigen Kapiteln gezeigt, reichen diese von einfachen statistischen Methoden über Ansätze des klassischen ML bis hin zu komplexen neuronalen Netzen aus dem Bereich des DL.

Welche Methoden sich für einen konkreten Anwendungsfall eignen hängt stark von den Eigenschaften und der Anzahl der verfügbaren Daten ab. Anforderungen an die Interpretierbarkeit und Rechenaufwand des Modelles, sowie die Komplexität des überwachten Prozesses beeinflussen die Wahl weiterhin.

Auch in dieser Arbeit könnten verschiedene Modelle zur Anwendung kommen. Methoden wie der Autoencoder, Isolation Forest oder probabilistische Modelle sind denkbar. Ein Vergleich der verschiedenen Ansätze wäre anhand definierter Performanzmetriken sinnvoll und könnte in zukünftigen Arbeitsschritten erfolgen.

Für diese Arbeit wird jedoch der Fokus klar auf die OCSVM gelegt. Die Wahl der Methode begründet sich aus der Robustheit gegenüber Ausreißer, ihrer guten Interpretierbarkeit im Vergleich zu Verfahren des DL und ihrer Anwendung in der Industrie. Der Vergleich mit anderen Verfahren ist jedoch trotzdem noch erforderlich, um die AD weiter zu optimieren.

Die OCSVM wird bereits erfolgreich in verschiedenen Bereichen zur AD eingesetzt. Dazu gehören zum Beispiel:

- Stromnetzindustrie[56]
- Stahlindustrie[57]
- Wasseraufbereitungssysteme[58]

[56] WANG; u.a., 2019
[57] HU; u.a., 2020
[58] INOUE; u.a., 2017

## 3.4 Mathematische Beschreibung ausgewählter Verfahren

In dieser Arbeit wird die OCSVM als eine Methode zur AD eingesetzt und im Folgenden näher erläutert. Die OCSVM stellt eine spezialisierte Variante der klassischen Support Vector Machine (SVM) dar. Während die SVM ursprünglich als ein Verfahren zur Klassifikation binärer oder mehrerer Klassen entwickelt wurde, handelt es sich bei der OCSVM dagegen, wie der Name schon andeutet, um ein Ein-Klassen Modell. Diese Klasse repräsentiert das „Normale“ im Kontext der AD.[59]

### 3.4.1 Maximal Margin Classifier

Die Basis der SVM bildet ein linearer Klassifikator, der „Maximal Margin Classifier“. Ziel dieser Methode ist es eine Trennebene (Hyperebene) zwischen 2 Klassen zu finden. Es kann jedoch unendliche viele Lösungen für diese Hyperebene geben. Deshalb wird die Ebene gesucht, die den Abstand zwischen den nächstbefindlichen Punkten beider Klassen und der Trennebene maximiert. [60]

In Abbildung 11 ist diese Methode in einem zweidimensionalen Raum veranschaulicht. Die nächstliegenden Punkte mit dem kleinsten Abstand zur Hyperebene werden als Support Vektoren bezeichnet. Der Abstand zwischen Hyperebene und Supportvektoren wird als Margin bezeichnet. Das Intervall zwischen den Support Vektoren beider Klassen wird häufig auch als „Straße“ o. „Margin Band“ bezeichnet und enthält keine Datenpunkte.[61]

Das Optimierungsproblem des Maximal Margin Classifer ist in Formel 10 dargestellt. Ziel ist es, unter allen möglichen durch w (Richtungsvektor) und b (Bias) definierten Hyperebenen jene zu bestimmen, die den Abstand zum nächstgelegenen Datenpunkten der Klassen maximiert.[62]

[59] Vgl. SHANG; u.a., 2016 S. 1042
[60] Vgl. FRANC; HLAVÁČ, 2003 S. 1986
[61] Vgl. BERWICK S. 4
[62] Vgl. SHANG; u.a., 2016 S. 1042

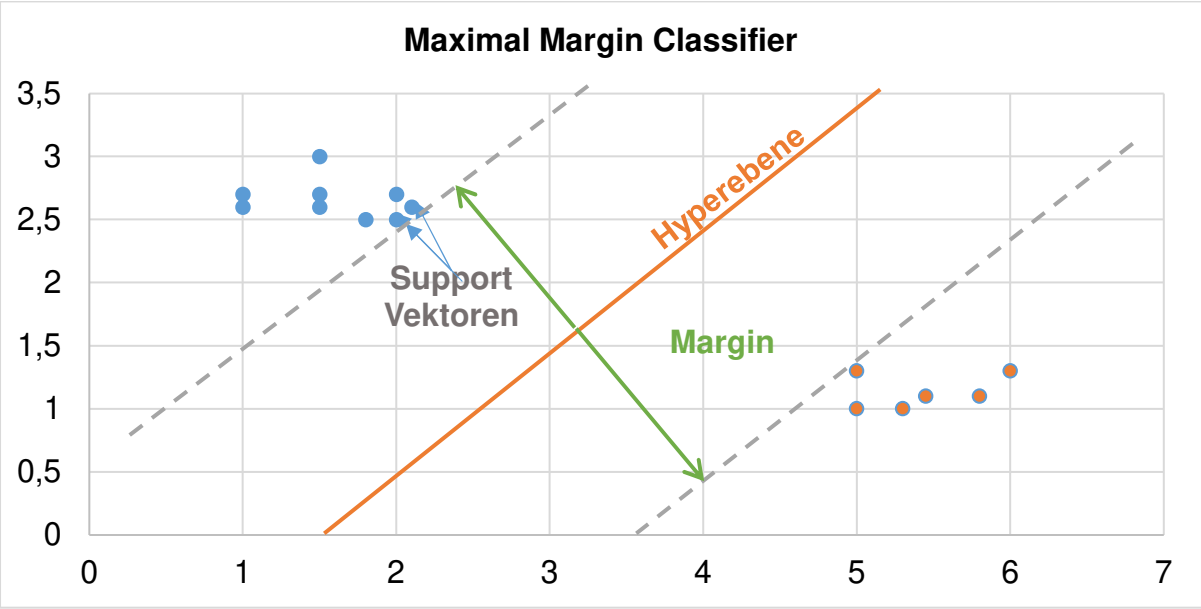



**Abbildung 11** Maximal Margin Classifier[63]

$$\underset{w,b}{argmax}\left[\min_{m=1,2,\dots,M} t_m\left(\frac{w^T * x_m + b}{\|w\|}\right)\right] = \underset{w,b}{argmax}\left[\frac{1}{\|w\|}\min_{m=1,2,\dots,M} t_m(w^T * x_m + b)\right]$$

**Formel 10** Optimierungsproblem des Maximal Margin Classifier [64]

Da jedoch eine Optimierung nach 2 Parametern schwierig ist wird die Funktion äquivalent umgeformt. Wenn $\pm k$ der kleinste Abstand der Hyperebene zu den Klassen ist und man $w$ und $b$ durch k teilt, erhält man ein äquivalentes Optimierungsproblem wie in Formel 11. Um das Maximum der Margin zu berechnen, benutzt man statt dem Maximum von $\frac{1}{\|w\|}$ das Minimum von $\frac{1}{2}\|w\|^2$.

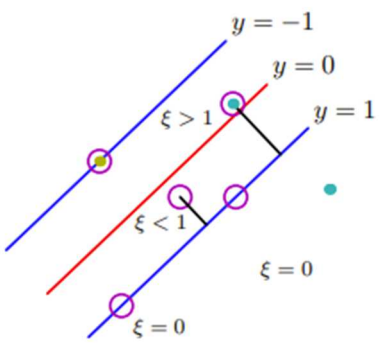
$$\underset{w,b}{argmin}\left(\frac{1}{2}\|w\|^2\right) \mid t_m(w^T * x_m + b) \geq 1, m = 1,2,\dots,M$$

**Formel 11** Umgestelltes Optimierungsproblem des Maximal Margin Classifer[65]

Der Maximal Margin Classifer ist jedoch anfällig gegenüber Ausreißern. Einzelne falsch liegende Datenpunkte können den Klassifizierer stark beeinflussen und somit zu Fehlklassifikationen bei der Prädiktion von neuen Betrachtungen führen. Zudem können reale Datensätze oft nicht linear voneinander getrennt werden.[66]

[63] Eigene Abbildung vgl. POENISCH S. 9
[64] FRANC; HLAVÁČ, 2003 S. 1986
[65] POENISCH S. 11.f
[66] Vgl. BISHOP, 2006 S. 331 f.

### 3.4.2 Support Vector Classifier




Der Support Vector Classifier wurde eingeführt, um dieses Problem zu lösen und nutzt das Konzept der Soft Margin. Dabei werden Falschklassifikationen durch eine Slackvariable $\xi$ erlaubt. (vgl. Abbildung 12) Gleichzeitig wird der erste anpassbare Parameter $C$ zur Regularisierung (Hyperparameter) eingeführt. Je größer C, desto höher ist die Bestrafung von Ausreißern. Er reguliert also die Menge der zugelassenen Fehlklassifikationen. [67]

**Abbildung 12** Slackvariable $\xi$ [68]

Das Optimierungsproblem mit einer Soft Margin ist in Formel 12 dargestellt. Abbildung 13 zeigt den Unterschied zwischen dem Maximal Margin Classifier (rot) und dem Support Vector Classifer (grün) im Fall von vorhandenen Ausreißern.

$$\underset{w,b}{argmin}\left(\frac{1}{2}\|w\|^2 + C * \sum_{m=1}^{M} \xi_m\right) \mid t_m(w^T * x_m + b) + \xi_m \geq 1, \xi_m \geq 0, m = 1,2,\dots,M$$

**Formel 12** Optimierungsproblem Support Vector Classifer mit Soft Margin[69]

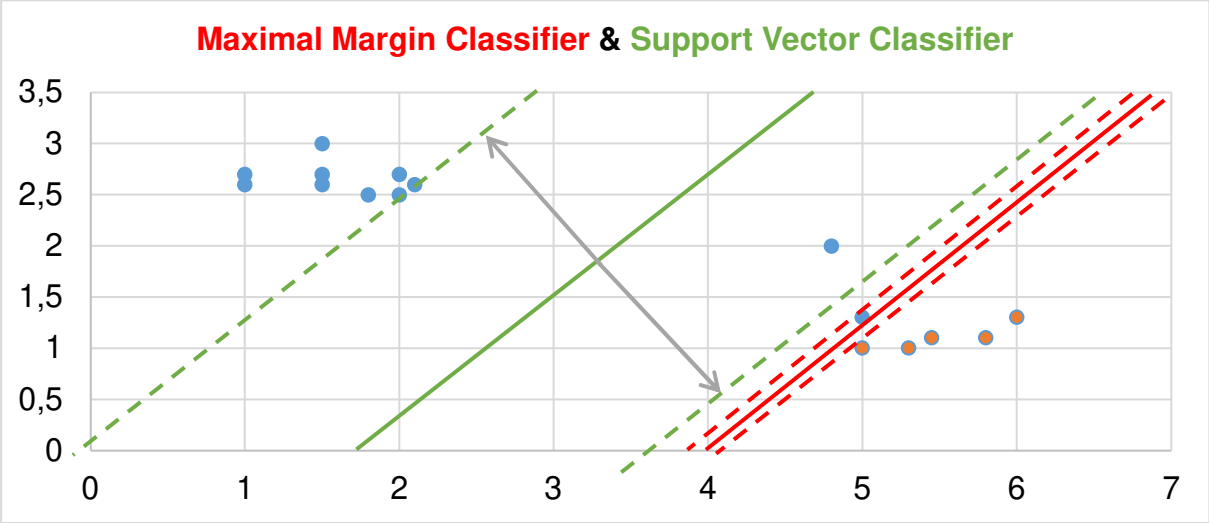



**Abbildung 13** Maximal Margin Classifier und Support Vector Classifer bei Daten mit Ausreißern[70]

[67] Vgl. POENISCH S. 20
[68] BISHOP, 2006 S. 332
[69] Vgl. POENISCH S. 20
[70] Eigene Abbildung

In vielen Fällen kann es jedoch sein, dass die Klassen nicht linear voneinander trennbar sind. In einem eindimensionalen Raum können zwei Klassen, z.B. wie in Abbildung 14, verteilt sein. Die Daten sind nicht mit einer Hyperebene trennbar.[71] Eine einfache Lösung für dieses Problem wäre es das Merkmal x auf einen Wert x^2 abzubilden, wodurch sich die Klassen im erweiterten Raum linear separieren lassen.

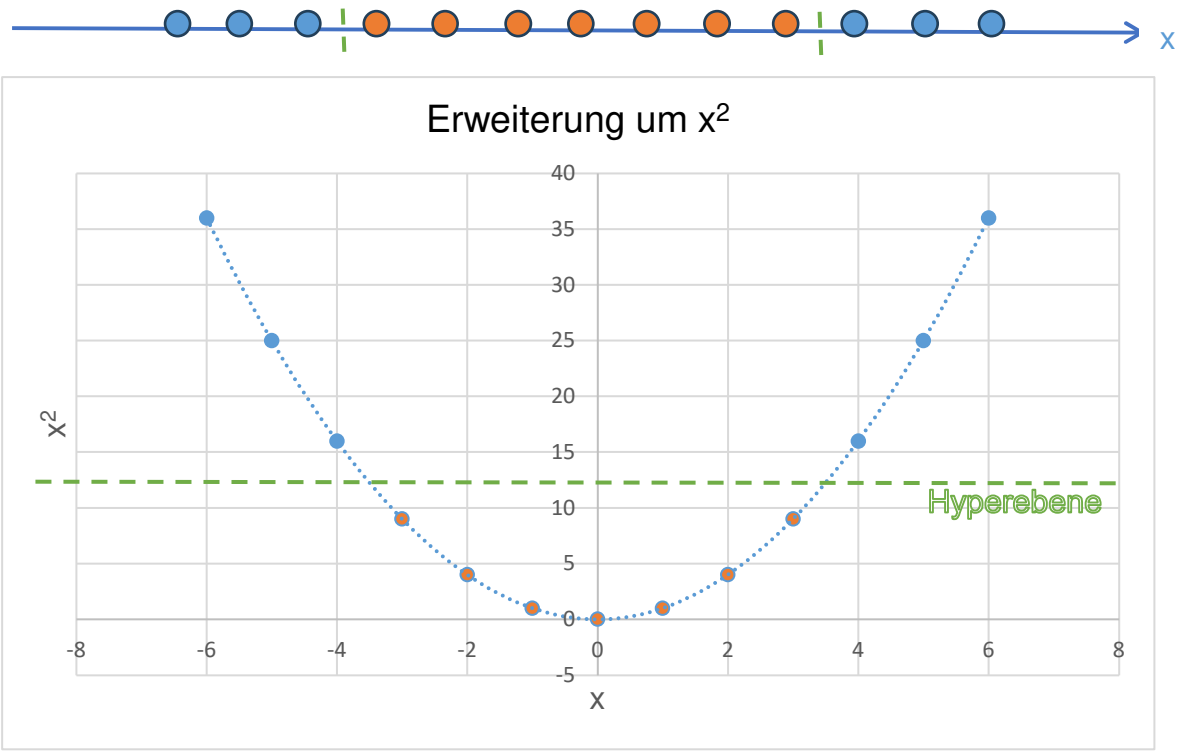



**Abbildung 14** Lineare Trennung im höherdimensionalen Raum[72]

### 3.4.3 Support Vector Machine

Wird der Support Vector Classifier um sogenannten Kernel erweitert spricht man von einer SVM. Kernel sind Funktionen, welche den Eingabevektorraum in eine höhere Dimension heben. Aus dem Eingabevektor x wird ein neuer Vektor $\phi(x)$ abgebildet, welcher bis zu unendlich Dimensionen groß ist. Durch geeignete Kernelfunktionen $k(x_i, x_k)$ kann das Skalarprodukt zwischen den transformierten Eingaberäumen $\phi(x)$ berechnet werden, ohne den höherdimensionalen Raum selbst berechnen zu müssen.[73] (vgl. Tabelle 1)

[71] Vgl. FRANC; HLAVÁČ, 2003 S. 1986
[72] Eigene Abbildung
[73] Vgl. BISHOP, 2006 328 f.

| (1) $x \rightarrow \phi(x)$ | |
|---|---|
| $\Re^N \rightarrow \Re^\infty$ | |
| (2) $\phi(x_i)^T * \phi(x_k)$ | (3) $k(x_i, x_k)$ |
| $\Re^N x \Re^N \rightarrow \Re^\infty x \Re^\infty \rightarrow \Re$ | $\Re^N x \Re^N \rightarrow \Re$ |

**Tabelle 1** Abbildungen des Eingabevektors in höherdimensionalen Raum[74]

(1) …Abbildung des Eingabevektors in unendlich dimensionalen Raum

(2) …Skalarprodukt des unendlich dimensionalen Raumes

(3) …Kernelfunktion

Beispielhafte Kernelfunktionen sind:

$$k(x, y) = x^T * y$$

**Formel 13** Linear Kernel ohne Hyperparameter

$$k(x, y) = (\mathrm{x}^\mathrm{T} * y + r)^d$$

**Formel 14** Polynomial Kernel[75] mit Hyperparametern $r, d$

$$k(x, y) = e^{-\gamma \|x-y\|^2}$$

**Formel 15** Radial Base Function Kernel[76] mit Hyperparameter $\gamma$

### 3.4.4 One Class Support Vector Machine

Die OCSVM stellt ein Modell dar, um zwischen normalen Verhalten und abweichenden Mustern zu unterscheiden. Dieses Verfahren basiert auf dem Konzept der SVM. Dafür lernt die OCSVM die Klasse des Normalen an. Die OCSVM überträgt das Konzept der SVM auf den Spezialfall der Ein-Klassen Klassifikation. Statt zwei Klassen voneinander zu trennen, versucht die OCSVM die „normalen“ Daten vom Ursprung des Raumes möglichst kompakt zu trennen. Das Optimierungsproblem ist in Formel 16 dargestellt.[77]

$$\min_{w,\rho,\xi} \frac{1}{2}||w||^2 + \frac{1}{v * n}\sum_{i=1}^{n} \xi_i - \rho \mid w^T\phi(x_i) \geq \rho - \xi_i\,, \qquad \xi_i \geq 0$$

**Formel 16** Primal Problem der OCSVM[78]

[74] Vgl. POENISCH S. 17
[75] Vgl. online: KUNZE, 2004 (18.07.2025) S. 16
[76] Vgl. online: KUNZE, 2004 (18.07.2025) S.15 f.
[77] Vgl SHANG; u.a., 2016 S.1042
[78] Vgl SHANG; u.a., 2016 S.1042

Hierbei gibt es mehrere mögliche Hyperparameter. Einerseits gibt es den Parameter $v$ („nu“). Dieser gibt an wie viele Support Vektoren zur Bestimmung der Hyperebene mindestens benutzt werden. Gleichzeitig ist er ein Maß dafür wie viel Prozent der Beobachtungen des Trainingsdatensatzes maximal als Ausreißer klassifiziert werden dürfen.[79]

Andererseits gibt es die Hyperparameter der Kernel Funktionen. Im Falle des RBF-Kernels gibt es somit einen weiteren Parameter $\gamma$. Die Wahl des Kernels sowie die Wahl geeigneter Hyperparameter ist entscheidend für die Erkennungsleistung von Anomalien.

$$L_a(w,\rho,\xi,\alpha,\beta) = \frac{1}{2}||w||^2 + \frac{1}{v*n}\sum_{i=1}^{n}\xi_i - \rho - \sum_{i=1}^{n}(w^T\phi(x_i) - \rho + \xi_i)\alpha_i - \sum_{i=1}^{n}\xi_i\beta_i$$

**Formel 17** Lagrange Funktion der OCSVM[80]

Die Lagrange Funktion wird in ein duales Problem umgestellt, in dem alle partiellen Ableitungen der Funktion gleich 0 gesetzt werden. Diese werden umgestellt und entweder in die Ausgangsfunktion eingesetzt oder sie ergeben eine Nebenbedingung. Daraus ergibt sich Formel 18, in der zusätzlich das Skalarprodukt durch die Kernelfunktion ersetzt wird.

$$\min_{\alpha} \sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_i\alpha_j K(x_i,\, x_j) \;\; | \; 0 \le \alpha_i \le \frac{1}{v*n}, \qquad \sum_{i=1}^{n}\alpha_i \; = 1$$

**Formel 18** Duales Problem der OCSVM[81]

Weiterhin ergibt sich die Entscheidungsfunktion aus allen Support Vektoren. Support Vektoren sind alle Punkte, für die $\alpha_i > 0$ gilt. Aus den Supportvektoren und dem Offset $\rho$ ergibt sich die Hyperebene. Der Abstand eines Punktes zur Hyperebene bestimmt seine Klasse. Ist der Abstand positiv gilt der Datenpunkt o. Vektor als normal, ist er dagegen negativ gilt er als Anomalie. (vgl. Formel 19)

$$f(x) = sign\left(\sum_{i=1}^{n}\alpha_i K(x_i,x) - \rho\right) \mid \rho = \frac{1}{|S|}\sum_{k\in S}\sum_{j=1}^{n} a_i K\big(x_j, x_k\big)$$

**Formel 19** Entscheidungsfunktion der OCSVM[82]

Die OCSVM trifft somit eine Entscheidung auf Basis eines geometrischen Kriteriums, dass sich aus dem angelernten Modell ergibt.

[79] Vgl. GOULEBE, 2024 S. 11
[80] WANG; u.a., 2019 S. 181582
[81] WANG; u.a., 2019 S. 181582
[82] WANG; u.a., 2019 S. 181582

Auf Grundlage der in diesem Kapitel gelernten Konzepte wird im folgenden Abschnitt ein datengetriebenes AD-Modell auf Basis der OCSVM mit realen Prozessdaten trainiert und evaluiert.

# 4 Entwicklung eines datengetriebenen Anomalie-Erkennungssystems

## 4.1 Vorgehen nach dem CRISP-DM Prozess

In diesem Kapitel wird ein ML-Modell zur AD nach dem CRISP-DM-Prozess entwickelt. Der CRISP-DM-Prozess ist ein viel genutzter Standardprozess und durchläuft 6 Prozessphasen, mit iterativen Rückkopplungen, wie in Abbildung 15 dargestellt.[83]

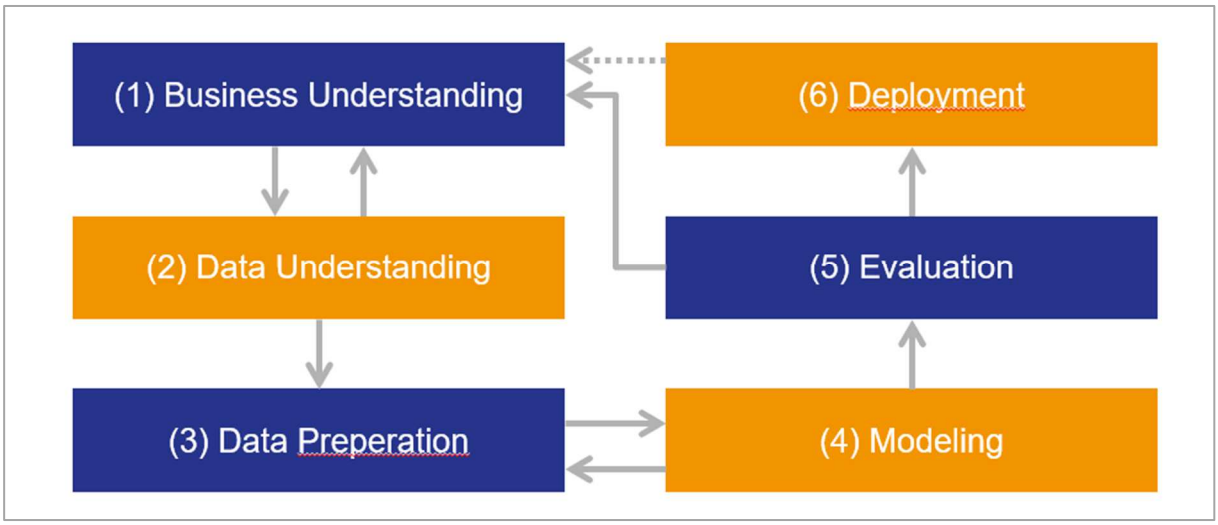



**Abbildung 15** CRISP-DM-Prozess [84]

Im Folgenden werden dafür die Prozessphasen 2-5 am Beispiel des Modells der OCSVM durchgeführt. Prozessphase 1 wurde bereits in Kapitel 2 durchgeführt.

## 4.2 Datensammlung und -beschreibung

Die zweite Phase des CRISP-DM-Prozesses umfasst die Sammlung und Beschreibung von Daten. Erste Analysen können hierbei schon Erkenntnisse zur Lösung des Problems liefern. [85]

Ziel dieser Arbeit ist es Prozessanomalien der Pulverförderung binär zu klassifizieren. Aus dem Produktionsprozess wird eine Vielzahl an Größen mit Sensoren gemessen. Für die Überwachung des Pulverflusses stehen dafür 3 von Sensoren erfassten Größen zur Verfügung. Diese werden in dieser Arbeit anonymisiert behandelt, normalisiert dargestellt und mit den Namen Merkmal 1-3 betitelt. Die Daten stammen

[83] Vgl. HURST; WENTZIEN; SCHMALZRIED, 2024 S. 252 f.
[84] Eigene Abbildung vgl. HURST; WENTZIEN; SCHMALZRIED, 2024 S. 252 f.
[85] Vgl. HURST; WENTZIEN; SCHMALZRIED, 2024 S. 253

hierbei aus einer Datenbank, welche die historischen Daten des Produktionsprozesses enthalten. Der erste Schritt der Analyse ist eine Darstellung der Merkmale nach der Zeit t (Zeitreihe). Ein Ausschnitt der Zeitreihen ist in Abbildung 16 visualisiert.

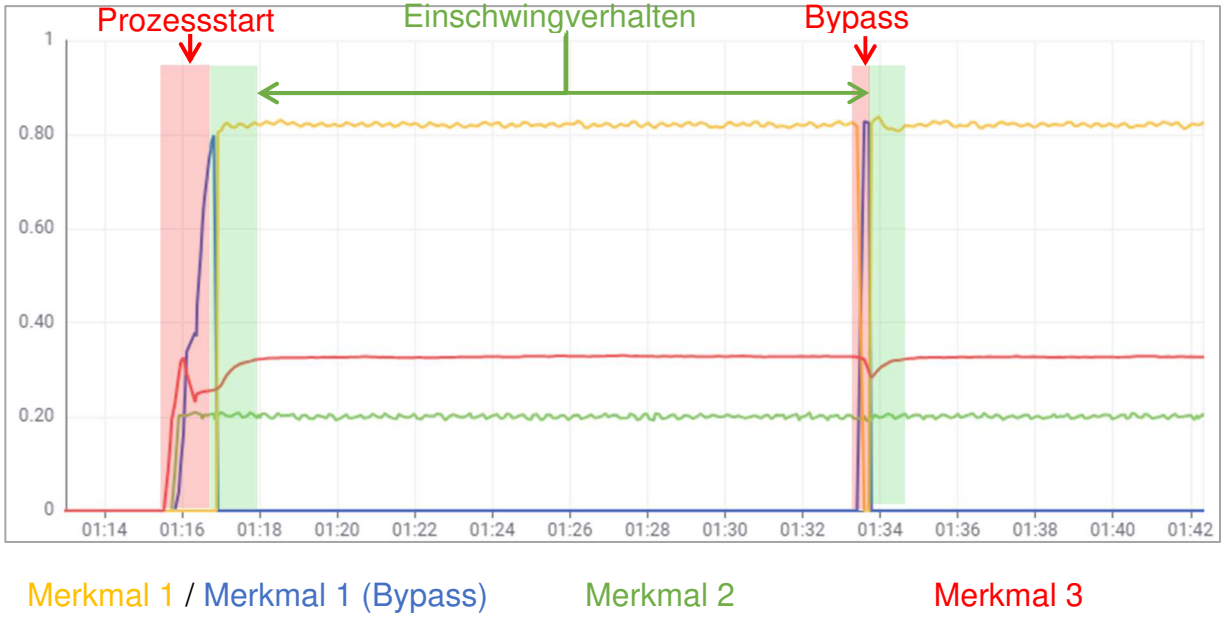



**Abbildung 16** Zeitreihenausschnitt der Pulverlinienmerkmale[86]

Je nach Ventilstellung wird das Pulver vom Pulvertopf zum Brenner oder in einen Sammelbehälter geleitet. Durch unterschiedliche Leitungseigenschaften tritt, bei der Umstellung des Ventils, dynamisches Verhalten auf. Wird im Folgenden von einem Bypass gesprochen, dann ist dies genau dann der Fall, wenn das Pulver in den Sammelbehälter fließt.

Weiterhin kann man unterscheiden, ob der Prozess aktiv oder inaktiv ist. Hierbei sind für die weitere Analyse nur die Zeiträume interessant, in denen der Prozess aktiv ist und das Pulver zum Brenner geleitet wird. $(\text{Merkmal } 1 \neq 0)$

Der Beschichtungsprozess kann in verschiedene Phasen eingeteilt werden:

1) Prozess inaktiv
2) Prozessstart
3) Bypass
4) Einschwingverhalten durch Bypassschaltung
5) Stationäres Verhalten

Bei der Förderung des Pulvers zum Brenner kommt es durch den Bypass zur ungleichmäßigen Verteilung zwischen dem Einschwingverhalten und stationären Verhalten. Aufgrund des Ungleichgewichtes und den Abweichungen der Datenpunkte im

[86] Eigene Abbildung

dynamischen im Vergleich zum stationären Verhalten kann es sein, dass Modelle diese Bereiche häufiger als Anomalie klassifizieren.

Außerdem kann es sein, dass der für das Einschwingverhalten typische Verlauf für ein Merkmal auftritt, jedoch keine Bypassschaltung unmittelbar vorher stattgefunden hat. Es kann somit sein, dass dieses Auftreten eine kontextuelle Prozessanomalie darstellt. (vgl. Abbildung 17)

Merkmal 1 Merkmal 2 Merkmal 3

**Abbildung 17** Kontextuelle Anomalie im Kontext der Bypassschaltung[87]

Weiterhin werden die Datenpunkte in nicht äquidistanten Abständen, zwischen einer und fünf Sekunden aufgezeichnet. Dies kann zum Beispiel zu Problemen bei der Detektion zeitlich kontextueller Anomalien und mit Analyseverfahren führen. Dies muss bei der weiteren Analyse berücksichtigt werden, indem zum Beispiel fehlende Werte ergänzt werden. [88]

Hierbei handelt es sich um einen Datenvorverarbeitungsschritte, welche Teil der dritten CRISP-DM-Prozessphase sind. Die Datenaufbereitung wird im folgenden Kapitel beschrieben.

[87] Eigene Abbildung
[88] Vgl. STOETZER, 2020 S. 111

## 4.3 Vorverarbeitung: Bereinigung, Normalisierung und Feature Engineering

Ziel der dritten Phase ist es die Daten in eine geeignete Struktur dem Modell zu übergeben. Dazu gehört es die Daten zu selektieren, bereinigen, manipulieren, formatieren und zu integrieren.[89] Abbildung 18 zeigt den Datenvorverarbeitungsprozess für die AD auf der Pulverlinie.

Für den Trainings- oder Prädiktionszeitraum werden die Daten aus dem Datenbanksystem abgefragt und stehen wie in Tabelle 2 zur Verfügung.

| Attribut | Zeitpunkt t | ID-Brenner | Stellung Bypass | Merkmal 1 | Merkmal 2 | Merkmal 3 |
|---|---|---|---|---|---|---|
| Datentyp | Date Time | Integer | Bool | Double | Double | Double |

**Tabelle 2** Attribute und Datentypen des Datensatzes

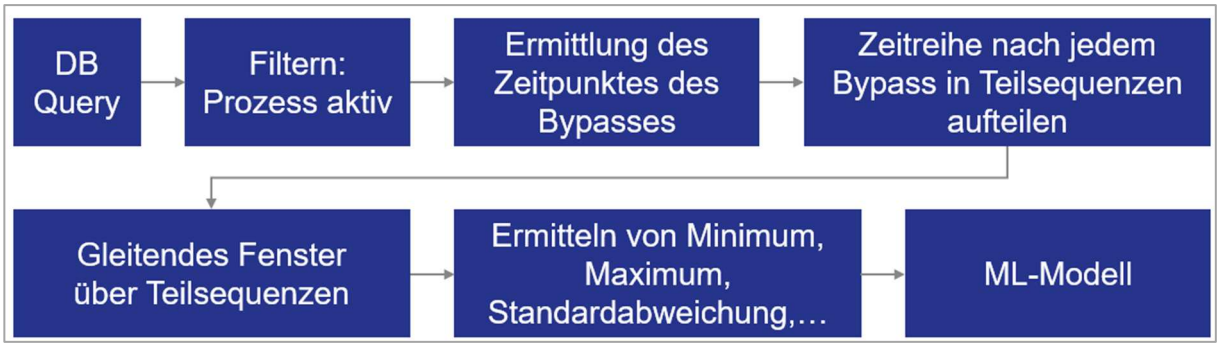



**Abbildung 18** Datenvorverarbeitungsprozess[90]

Da nur die Datenpunkte interessant sind, in denen der Prozess aktiv ist, werden zunächst nur die Datenpunkte gebraucht bei denen $\text{Merkmal } 1 \neq 0$ ist.

Weiterhin sind nur die Datenpunkte interessant an denen die Bypassschaltung inaktiv ist. Aufgrund des Einschwingverhaltens nach der Bypassschaltung muss jedoch eine Information darüber existieren, wann diese zuletzt stattgefunden hat. Um die Trennung der Datenpunkte je nach Ventilstellung des Bypasses zu gewährleisten und trotzdem den Kontext dieser Schaltung beizubehalten muss diese Information übertragen werden.

Dazu wird der erste Datenpunkt ermittelt, welcher nach einem Wechsel des Bypass-Parameters von aktiv auf inaktiv auftritt. Dafür wird in der zeitlich geordneten Relation

[89] Vgl. REUß; ZWIESLER S.
[90] Eigene Abbildung

jeder Parameterwert über die Ventilstellung des Bypasses an die nächste Reihe übergeben. (vgl. Tabelle 3) Anschließend werden alle Datenpunkte gefiltert, bei denen der Bypass inaktiv ist oder Null Werte existieren. (vgl. Tabelle 4)

| Stellung Bypass | Stellung Bypass (-1) | … |
|---|---|---|
| 0 | Null | … |
| 1 | 0 | … |
| 1 | 1 | … |
| 0 | 1 | … |
| 0 | 0 | … |

**Tabelle 3** Verzögerungsspalte der Bypassstellung

| Stellung Bypass | Stellung Bypass (-1) | … |
|---|---|---|
| 0 | 1 | … |
| 0 | 0 | … |

**Tabelle 4** Filterung der Datenpunkte mit inaktivem Bypass

In Abbildung 19 ist die Zeitreihe nach der Transformation des Datensatzes dargestellt. Im nächsten Schritt wird der Datensatz in Teilsequenzen aufgeteilt, welche nach einer Bypassschaltung beginnen und vor der nächsten enden.

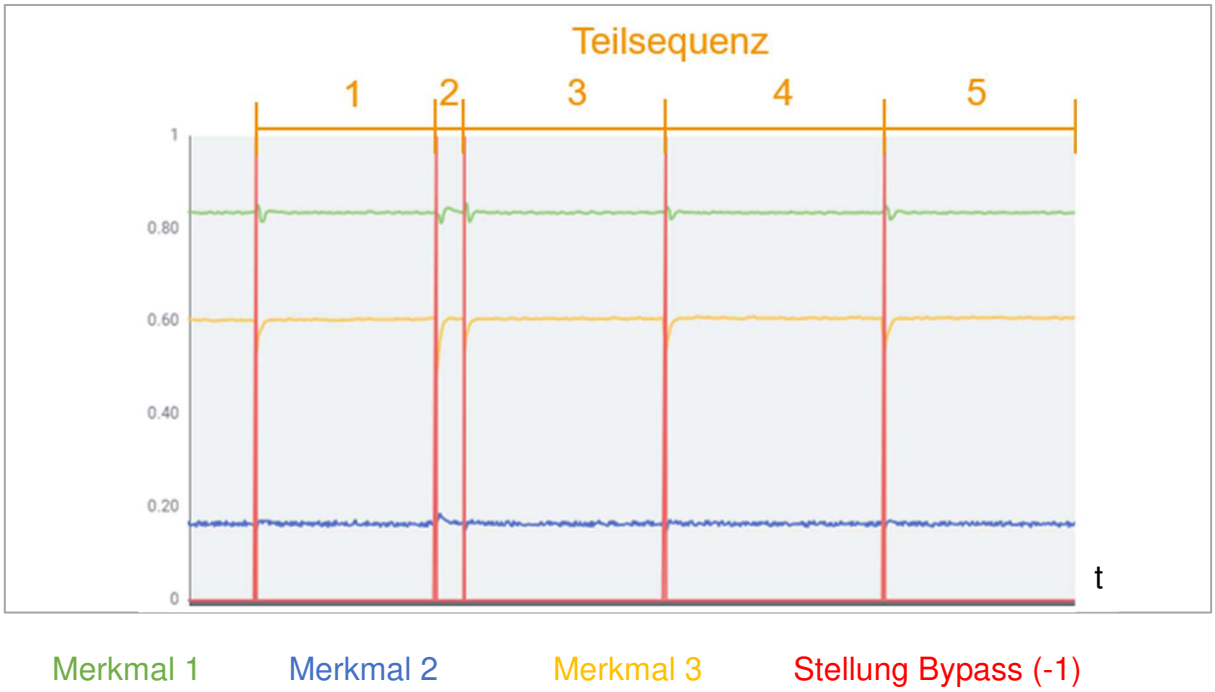



**Abbildung 19** Transformierte Daten und Einteilung in Teilsequenzen[91]

Weiterhin wird jede Teilsequenz mit der Methode des gleitenden Fensters in Stichproben eingeteilt.[92] Dafür wird ein Fenster der Länge von T Zeiteinheiten

[91] Eigene Abbildung
[92] Vgl. RAFFERTY; u.a., 2016 S. 2539

beginnend von dem ersten Datenpunkt der Teilsequenz von links nach rechts mit einer Schrittweite von s Zeiteinheiten geschoben. (vgl. Abbildung 20)

Es erfolgt weiterhin die Ermittlung der Fensteranzahl seit der letzten Bypassschaltung. So bleibt der Kontext der Bypassschaltung erhalten und kann dem Modell mit übergeben werden.

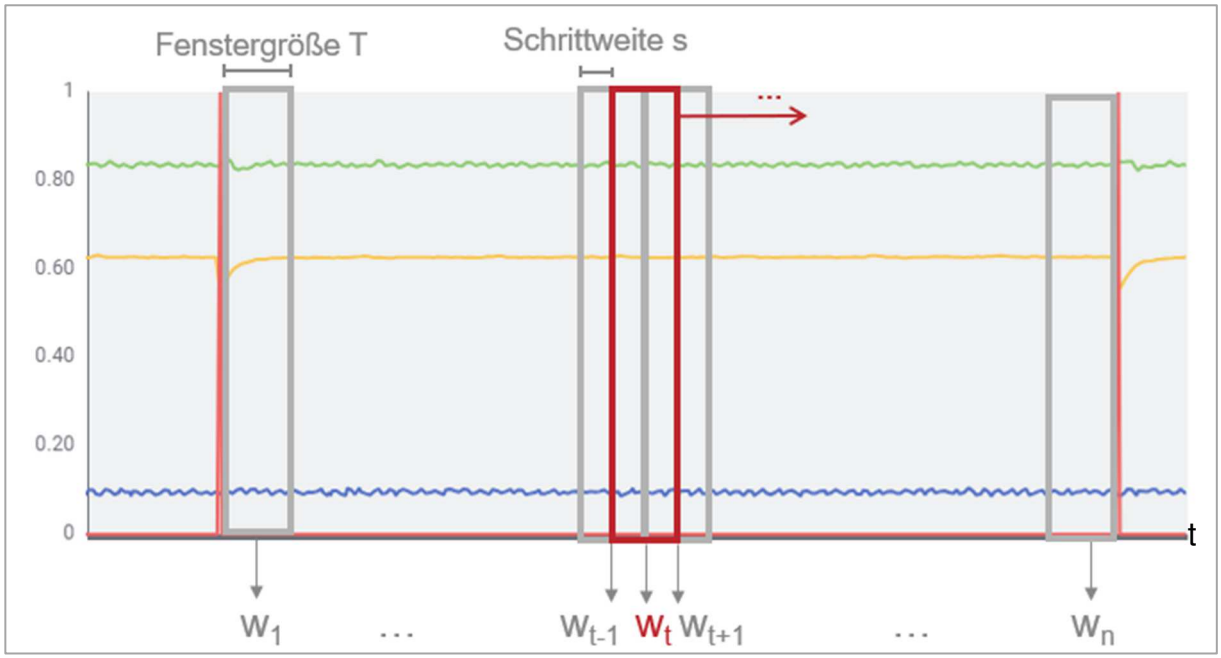



**Abbildung 20** Methode des gleitenden Fensters[93]

Jede Stichprobe hat hierbei durch die zeitlich nicht äquidistanten Abstände eine unterschiedliche Anzahl an Datenpunkten. Es muss garantiert werden, dass jedes Fenster gleich viele Dimensionen besitzt. Um die Datenpunkte in einer Stichprobe in ein einheitliches Format zu bringen, müssen die Datenpunkte zusammengefasst oder erweitert werden.

Dazu können Datenpunkte zum Beispiel interpoliert und somit die Anzahl der Datenpunkte erhöht (Upsampling) oder die Anzahl der Datenpunkte, durch Quantisierung, verringert werden. (Downsampling)

Bei der Interpolation werden die fehlende Datenpunkte durch geschätzte Werte, um die regelmäßigen Abstände herzustellen.[94] Die originalen Datenpunkte bleiben hierbei erhalten. Es findet also kein Informationsverlust statt. Jedoch erhöhen sich die benötigte Rechenleistung und der Speicherplatz für die AD.

[93] Eigene Abbildung
[94] Vgl. RUNKLER, 2015 S. 26

Bei der Quantisierung wird die Anzahl der Datenpunkte verringert. Benötigter Speicherplatz und Rechenleistung können so reduziert werden. Je nach Vorgehen kommt es hierbei zu mehr oder weniger Informationsverlust.[95]

Beide Varianten bringen die Daten in ein einheitliches Format. Aufgrund der Verringerung des Speicherbedarfes und der Rechenzeit, werden die Stichproben in ihrer Dimension reduziert. Zudem werden die Daten durch diesen Schritt verallgemeinert. Um dies zu ermöglichen, werden für jedes Fenster statistische Eigenschaften von der Stichprobe ermittelt. Zu den ermittelten Eigenschaften gehören:

- Mittelwert
- Median
- Standardabweichung
- Min
- Max
- Schiefe
- Wölbung

Durch diesen Schritt wird die Anzahl der Beobachtungen von 398697 mit 3 Dimensionen auf 14832 mit 21 Dimensionen reduziert. Dadurch reduziert sich die Größe des Datensatzes um 73.96 %. (vgl. Formel 20)

$$\frac{\mathrm{M} * \mathrm{W}}{\mathrm{N}} * 100 - 100 = \frac{7 * 14832}{398697} * 100 - 100 \approx -73{,}96\ \%$$

**Formel 20** Reduktion der Größe des Datensatzes durch Merkmalsextraktion

M…Anzahl der extrahierten Merkmale

W…Anzahl der Fenster

N…Anzahl der Datenpunkte

Jedoch können sich einige nominale Werte (Sollwerte) im Laufe der Zeit ändern oder es können Trendbewegungen auftreten. Wird ein Modell ausschließlich in einem Zeitraum trainiert, in dem nur ein fest definierter Sollwert vorliegt, dann führt jede Änderung dieser Konfiguration unmittelbar zu einer Abweichung zum trainierten Normalverhalten.

Um dies zu vermeiden, werden in dieser Arbeit nur die gegenüber dem Mittelwert verschobenen Kurven in die AD einbezogen. Die statistischen Eigenschaften, mit Ausnahme des Mittelwertes, werden aus diesem mittelwertzentrierten Fenstern extrahiert. Dieser Schritt verhindert, dass das Modell bei jeder Änderung der Führungsgrößen neu trainiert werden muss.

[95] Vgl. RUNKLER, 2015 S. 20

Die Mittelwerte der Kurve werden separat betrachtet, da auch deren Veränderung Hinweise auf Anomalien liefern können. Abgesehen vom Mittelwert verhalten sich die Kurven jedoch ähnlich, wodurch sich so Muster besser erkennen lassen.

Ohne diese Maßnahme würde es in den meisten Fällen sofort zu Fehlalarmen und Warnungen kommen. Sowohl das Training des Modells als auch die anschließende Bewertung von Fehlalarmen zeitaufwendige Maßnahmen. Zudem müssten für bei Änderungen der Führungsgrößenkonfigurationen zunächst neue Daten gesammelt werden. In dieser Phase ist das System nur eingeschränkt oder gar nicht nutzbar.

## 4.4 Training der One Class Support Vector Machine

### 4.4.1 Beschriftung der Prüfdaten

Auch wenn die OCSVM ein unüberwachtes Verfahren ist, ist für die Bewertung des Modells und dessen Optimierung eine beschriftete Prüfmenge nötig.[96]

Diese Daten enthalten sowohl normale als auch anormale Instanzen, welche gelabelt (beschriftet) sind. So lässt sich objektiv die Leistung des Modells bewerten, indem überprüft wird, wie gut das Modell echte Anomalien (richtig Positiv) oder normale Daten (richtig Negativ) korrekt klassifiziert. [97]

Deshalb wurde der gesamte Datensatz von 14832 Fenstern nach auffälligen Anomalien untersucht und die Fenster wie folgt beschriftet:

- 1 = „Normal“
- -1 = „Anormal“ / “Auffällig“

Dafür werden die Daten visuell in Form von Streudiagrammen (vgl. Abbildung 21) und die unverarbeiteten Zeitreihen (vgl. Abbildung 17) dargestellt. Wie in Abbildung 21 dargestellt werden über Entscheidungsgrenzen (rot=“univariat“, grün = „multivariat“) starke Ausreißer als Anormal beschriftet. Dabei wurde darauf geachtet, dass die Fenster, welche Einschwingverhalten durch die Regelung enthalten, getrennt betrachtet werden. Die Betrachtung hat aus Sicht verschiedener Parameter zueinander stattgefunden.

Weiterhin werden die Verlaufskurven der anormal markierten Fenster betrachtet und es wird nochmals überprüft, ob sie Anormal erscheinen. Auf dessen Grundlage folgt die iterative Anpassung der Beschriftungen.

[96] Vgl. RUFF, 2021 S. 6
[97] Vgl. RUFF, 2021 S. 12

Nach dem Beschriftungsprozess sind 237 von 14832 Fenstern als Anomalie gelabelt. Daraus folgt, dass 1,59 % des Datensatzes als Anomalie beschriftet sind. Der Beschriftungsprozess ist sehr Fehleranfällig. Es gibt also keine Garantie, dass alle Fenster richtig beschriftet sind.

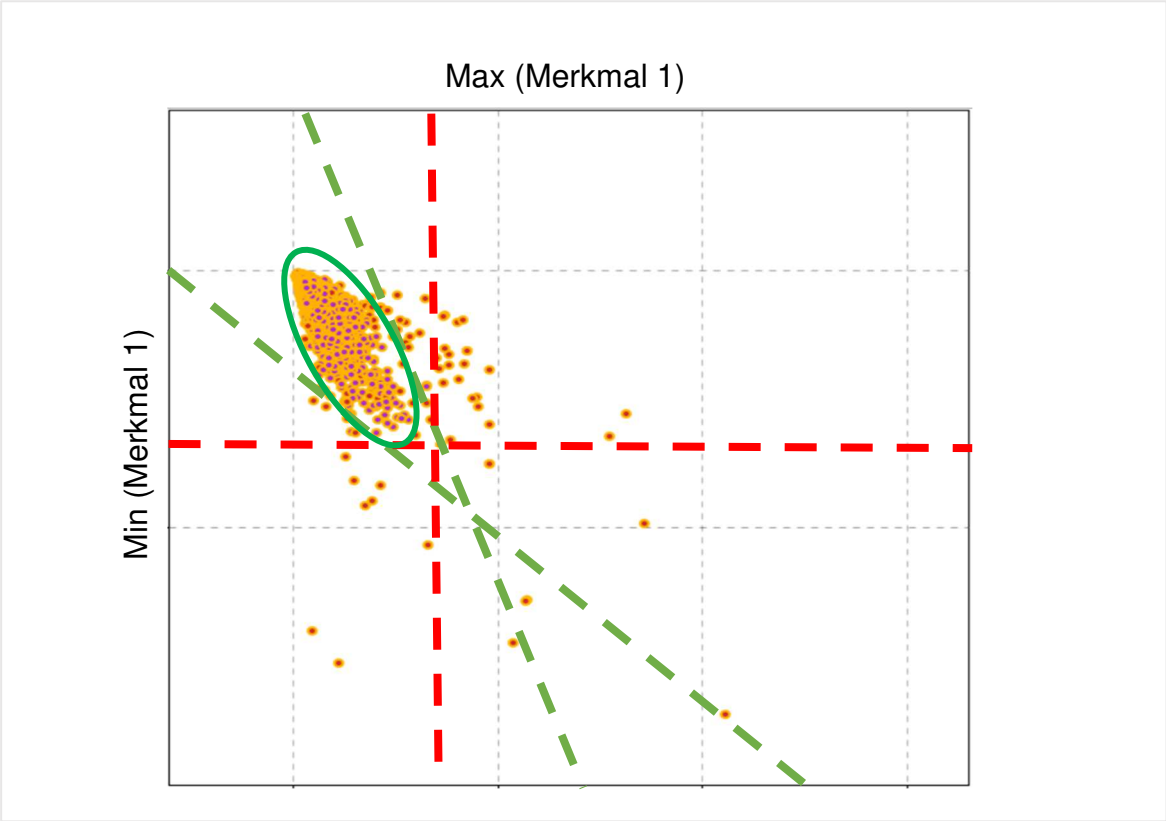



**Abbildung 21** Minimum-Maximum Streudiagramm der Stichproben von Merkmal 1 [98]

### 4.4.2 Metriken zur Bewertung der Prognosegüte

Der beschriftete Datensatz erlaubt es jedoch Metriken zu nutzen, um das Modell zu optimieren und zu evaluieren. Grundlage hierfür bildet die Konfusionsmatrix. (vgl. Tabelle 5) Die Komponenten der Konfusionsmatrix sind:

- Richtig Positiv (TP)
- Falsch Negativ (FN)
- Richtig Negativ (TN)
- Falsch Negativ (FN)

[98] Eigene Abbildung

| | | *Vom Modell prognostiziert* | |
|---|---|---|---|
| | | **Anomalie** | **Normal** |
| *Tatsächlicher Wert* | **Anomalie** | TP | FP |
| | **Normal** | FP | TN |

**Tabelle 5** Konfusionsmatrix im Kontext der Anomaliedetektion[99]

Aus diesen Elementen lassen sich die die folgenden Performancemaße (Formel 21-24[100] & 25) ableiten:

$$\frac{FN + FP}{FN + FP + TN + TP}$$

**Formel 21** Fehler

$$\frac{TP + TN}{FN + FP + TN + TP}$$

**Formel 22** Genauigkeit

$$\frac{TP}{TP + FP}$$

**Formel 23** Präzision

$$\frac{TP}{TP + FN}$$

**Formel 24** Sensitivität

$$\frac{(1 + \beta^2) * \text{Präzision} * \text{Sensitivität}}{\beta^2 * \text{Präzision} * \text{Sensitivität}}$$

**Formel 25** Fβ -Maß[101]

Der Fehler und die Genauigkeit sind ungeeignete Performancemaße, da im Falle der AD ein ungleichmäßig verteilter Datensatz vorliegt. Würde das Modell den gesamten Datensatz als „normal“ klassifizieren dann hätte es, bei 1% an Anomalien im Datensatz, eine Genauigkeit von 99 %.

Besser für ungleichgewichtet Klassen eignen sich die Sensitivität und die Präzision als Performancemaße. Die Präzision ist ein Maß dafür wie viele der als Anomalie prognostizierten Beobachtungen auch als Anomalie beschriftet wurden. Die Sensitivität beschreibt dagegen wie viele der als Anomalie gelabelten Daten vom Modell erkannt wurden.[102]

Um eine Balance zwischen der Sensitivität und der Präzision zu finden, kombiniert das Fβ -Maß beide Metriken. Dabei ist die Wahl von β entscheidend dafür, ob die Präzision

[99] Vgl. HUDE, 2020 S. 150
[100] HUDE, 2020 S 102, 151
[101] LEE; YANG; YOO, 2021 S. 2
[102] online: JAYASWAL, 2020 (24.07.2025)

oder die Sensitivität bevorzugt in dieses Maß eingeht. Ist β = 1, dann gehen beide Maße gleichermaßen ein. Anderenfalls gilt[103]:

- β > 1: Sensitivität stärker gewichtet
- 0 < β < 1: Präzision stärker gewichtet

Beim Beschriften der Anomalien im Datensatz kann es passiert sein, dass nicht alle Anomalien als Anomalie gelabelt wurden. Solche welche beschriftet wurden sind jedoch zum Großteil deutlich sichtbar und sollten vom Modell erkannt werden. Deshalb ist nach eigenem Ermessen ein höherer Wert der Sensitivität wichtiger. Für die weitere Hyperparameteroptimierung und Evaluation des Modells wurde sich folglich für das F2-Maß entschieden.

### 4.4.3 Hyperparameteroptimierung

Ziel ist es mit den beschrifteten Daten und unter Verwendung des F2-Maßes die Hyperparameter der OCSVM zu optimieren. Als Kernel wird die Radial-Base-Funktion festgelegt. Daraus ergeben sich zwei wesentliche Hyperparameter (vgl. Kapitel 3.3.4):

- „Nu“ $\nu \in (0, 1]$: Maximaler Anteil an Ausreißern in den Daten
- Gamma $\gamma > 0$: Parameter des RBF-Kernels

Da genau 2 Hyperparameter vorliegen, wird eine Gittersuche (Grid Search) über beide Parameter durchführt. Diese Methode testet systematisch alle möglichen Kombinationen an Hyperparameter aus einer definierten diskreten Teilmenge. (vgl. Abbildung 22) Anschließend wird jenes Modell mit dem besten Wert der gewählten Prüfmetrik ausgewählt. [104] In diesem Fall handelt es sich bei der Prüfmetrik um das F2-Maß.

[103] online: JAYASWAL, 2020 (24.07.2025)
[104] Vgl. LIASHCHYNSKYI; LIASHCHYNSKYI, 2019 S.3

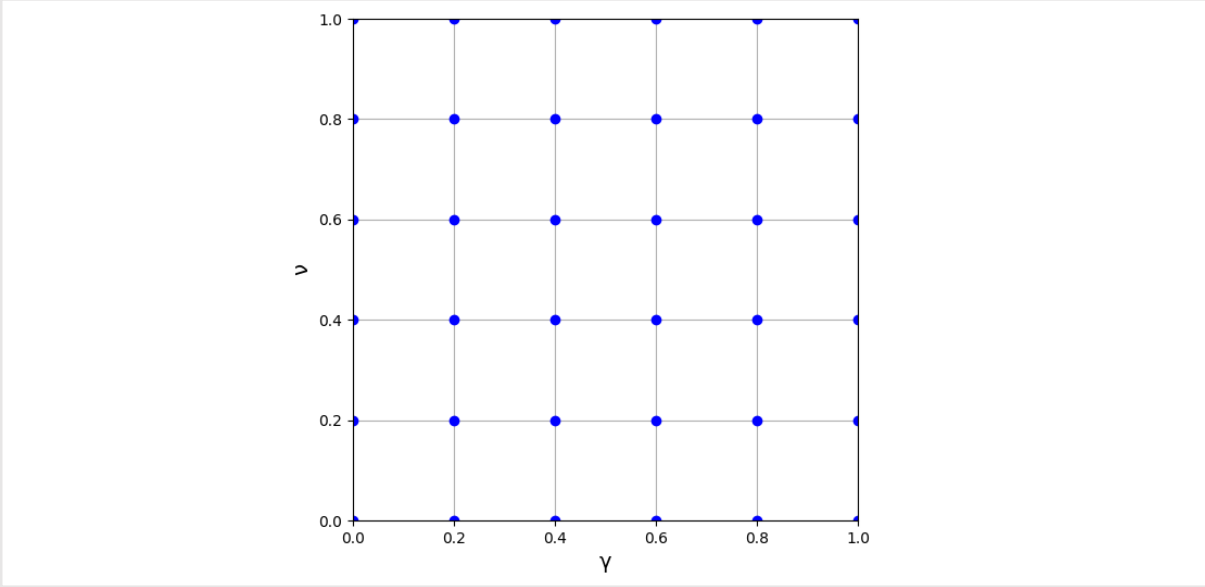



**Abbildung 22** Raum der Gittersuche[105]

Im Zuge der Gittersuche, wird zusätzlich zur zuverlässigen Bewertung der Modellleistung, Kreuzvalidierung eingesetzt. Dafür wird für jeden Satz an Hyperparametern ein k-fache Kreuzvalidierung durchgeführt, um Überanpassung zu vermeiden und robuste Parameter auszuwählen.

Dafür wird der Datensatz in gleich große Teilmengen (Faltungen) aufgeteilt. Die Teilmengen überschneiden sich hierbei nicht. In einem iterativen Prozess wird jeder dieser Teilmengen einzeln als Testmenge benutzt. (vgl. Abbildung 23) Alle anderen k - 1 Faltungen dienen als Trainingsdaten, wobei sich Trainingsdaten und Testdaten nicht überlappen. [106]

Damit jeder Datensatz gleich viele positive und negative Klassen enthält werden als Teilmengen stratifizierte Zufallsstichproben als Teilmengen genommen. Das heißt jede Teilmenge enthält prozentual gesehen ein fast identisches Verhältnis zwischen der Klasse Anomalie und der Klasse Normal, wie die Gesamtmenge. Dafür wird der Datensatz in die Schichten (Klassen) aufgeteilt. Anschließend werden die nach der Klasse getrennten Beobachtungen zufällig auf die Faltungen aufgeteilt.[107]

[105] Eigene Abbildung vgl. LIASHCHYNSKYI; LIASHCHYNSKYI, 2019 S.6
[106] Vgl. BERRAR, 2019 S. 5
[107] Vgl. online: Geschichtete bzw. stratifizierte Stichprobe | Lehrbuch Psychologie, 2025 (06.08.2025)

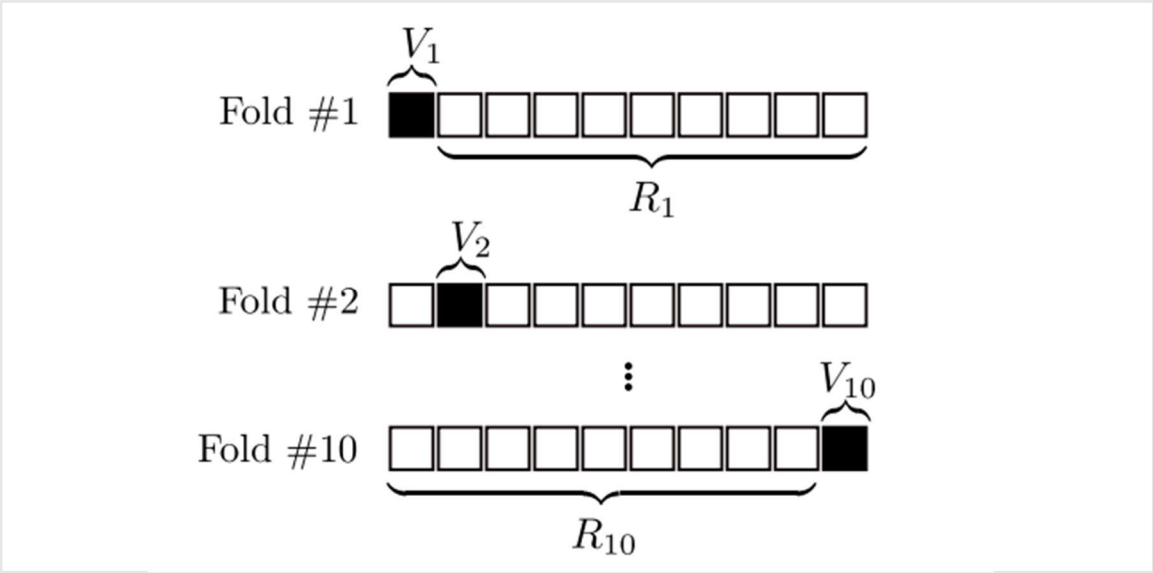



**Abbildung 23** Kreuzvalidierung mit k = 10 Teilmengen[108]

Die Hyperparameter werden mithilfe einer 5-fachen Kreuzvalidierung optimiert. Für die Gittersuche werden jeweils 10 Werte pro Hyperparameter als Teilmengen definiert. Die Suche verläuft dabei iterativ von grob nach fein:

Zunächst wird ein großer Bereich möglicher Hyperparameterkombinationen mit geringer Auflösung durchsucht, um eine grobe Orientierung zu erhalten. Anschließend wird der Wertebereich eingegrenzt, um eine feinere Auflösung zu erhalten.

Dieses Vorgehen ermöglicht eine effiziente Bestimmung der Hyperparameter, ohne unnötigen Rechenaufwand in weniger relevanten Bereichen des Raumes.

Dieser Prozess wird für folgende Varianten durchgeführt:

1) Training mit beschrifteten Anomalien + Merkmale standardisiert
2) Training mit beschrifteten Anomalien + Merkmale normalisiert von 0 bis 1
3) Training ohne beschriftete Anomalien + Merkmale normalisiert von 0 bis 1
4) Training ohne beschriftete Anomalien + Merkmale standardisiert

Die besten Hyperparameter und dazugehörigen Performancemaße sind in Tabelle 7 dargestellt.

| *Variante* | *Präzision* | *Sensitivität* | *F1-Maß* | *F2-Maß* | *Nu* | *Gamma* |
|---|---|---|---|---|---|---|
| *1* | 0.5925 | 0.7215 | 0.6503 | 0.6911 | 0.01733 | 0.023 |
| *2* | 0. 6023 | 0.7172 | 0.6527 | 0.6893 | 0.0139 | 2.65 |
| *3* | **0. 6814** | 0.7637 | **0.7192** | **0.7450** | 0.0023 | 1.9778 |
| *4* | 0.657 | **0.768** | 0.7081 | 0.7423 | 0.0028 | 0.023 |

**Tabelle 6** Ergebnisse der Gittersuche mit Kreuzvalidierung

[108] BERRAR, 2019 S. 5

Die schlechteste Performance hat die OCSVM hierbei bei Training mit Anomalien und normalisierten Merkmalen. Die beste Performance konnte dagegen mit normalisierten Daten und Training ohne Anomalien erreicht werden. In Abbildung 24 ist der dazugehörige Oberflächengraph der Gittersuche dargestellt.

Weiterhin fällt auf, dass der Parameter $\gamma$ (Gamma) stark davon abhängig ist, ob die Merkmale normalisiert oder standardisiert wurden. Werden sie normalisiert, dann ist Performance am besten, wenn ein größerer Wert für Gamma genutzt wird. Dies bedeutet, dass sich die Modellkomplexität erhöht. Dagegen ist ein geringer Gamma Wert besser, wenn die Daten standardisiert wurden.

Aufgrund der besseren Performance werden die Daten normalisiert und alle beschrifteten Anomalien aus dem Trainingsdatensatz entfernt.

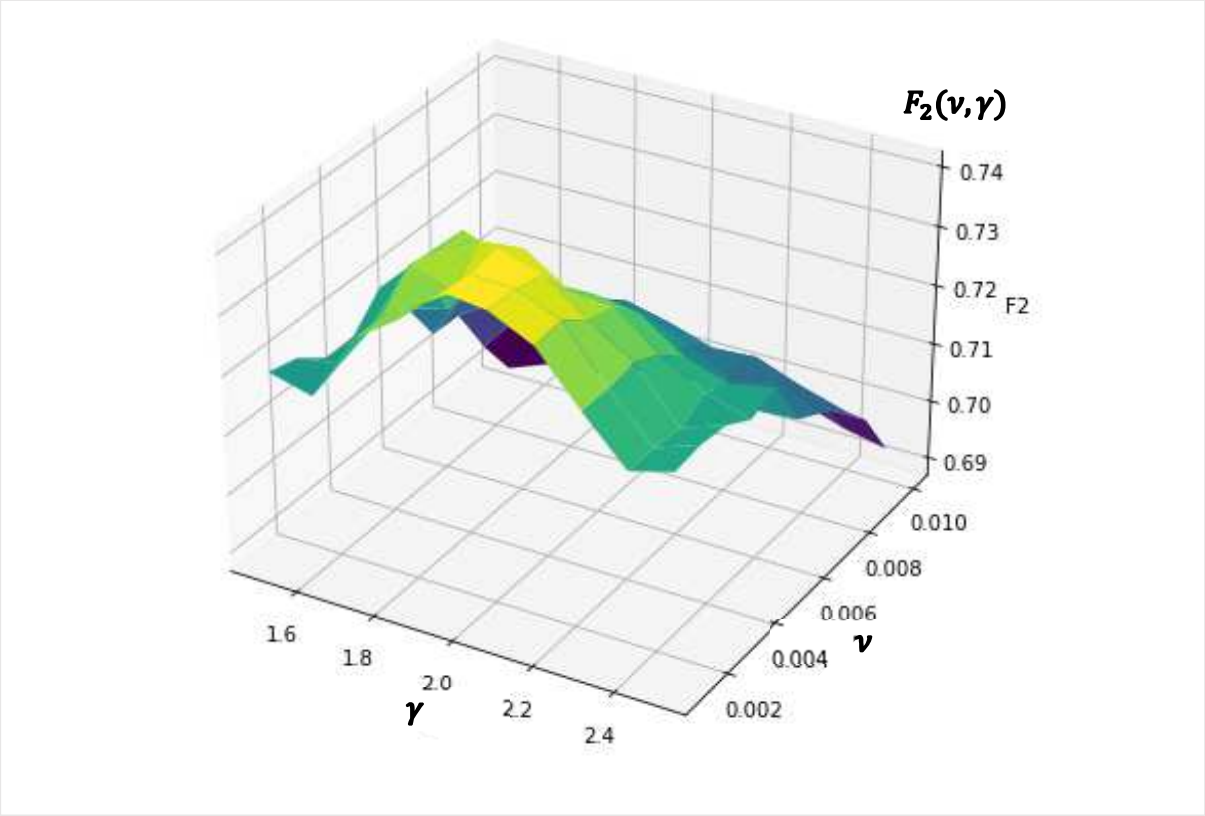



**Abbildung 24** Hyperparameteroptimierung von „nu“ und „gamma“ mittels Gittersuche[109]

[109] Eigene Abbildung

### 4.4.4 Evaluierung der Modellgüte

Nachdem das bestmögliche Modell über die Gittersuche ermittelt wurde, wird im Folgenden die Modellleistung näher betrachtet.

Wie in Kapitel 3.3.4 dargelegt grenzt die OCSVM die „normalen“ Daten mit einer Hyperebene ein. Jeden Eingabevektor des Modells kann man auch als einen Punkt im Raum verstehen. Dieser Punkt hat einen Abstand zur Hyperebene, auf dessen Grundlage entschieden wird, ob er zur Klasse „Normal“ oder „Anomalie“ gehört. Dies wird darüber entschieden, ob der Abstand positiv oder negativ ist. Man kann sich jedoch auch direkt den Abstand beziehungsweise den Entscheidungsfunktionswert vom Modell ausgeben lassen. Die Werte für alle Beobachtungen sind in Abbildung 25 in Form eines Histogramms dargestellt.

Dabei fällt auf, dass ein Großteil der vom Modell als Anomalie klassifizierten Daten (negative Distanz) einen großen Abstand zur Trennebene haben. Die Daten der normalen Klasse sind dagegen relativ normalverteilt. Nur relativ wenige Beobachtungen sind hierbei nahe der Entscheidungsgrenze. Die meisten Beobachtungen haben also einen großen Abstand zur Trennebene, wobei Anomalien eine deutliche Distanz zu dieser haben.

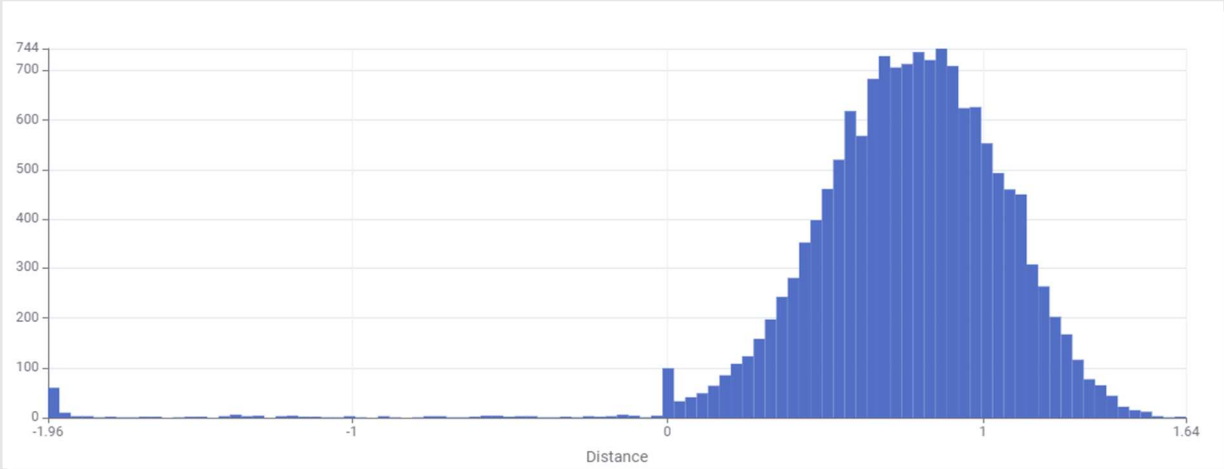



**Abbildung 25** Histogramm der Entscheidungsfunktionswerte[110]

Diese Betrachtung kann auch getrennt nach den Elementen der Konfusionsmatrix durchgeführt werden. Die Konfusionsmatrix der Trainingsdaten ist in Tabelle 7 dargestellt.

[110] Eigene Abbildung

Die Distanzen getrennt nach Klassen sind in Abbildung 26-28[111] dargestellt. Dabei fällt auf, dass es vor allem zu falsch positiven Klassifizierungen nahe der Entscheidungsgrenze kommt.

| | | *Vom Modell prognostiziert* | |
|---|---|---|---|
| | | **Anomalie** | **Normal** |
| *Tatsächlicher Wert* | **Anomalie** | 178 | 59 |
| | **Normal** | 46 | 14540 |

**Tabelle 7** Konfusionsmatrix der Trainingsergebnisse

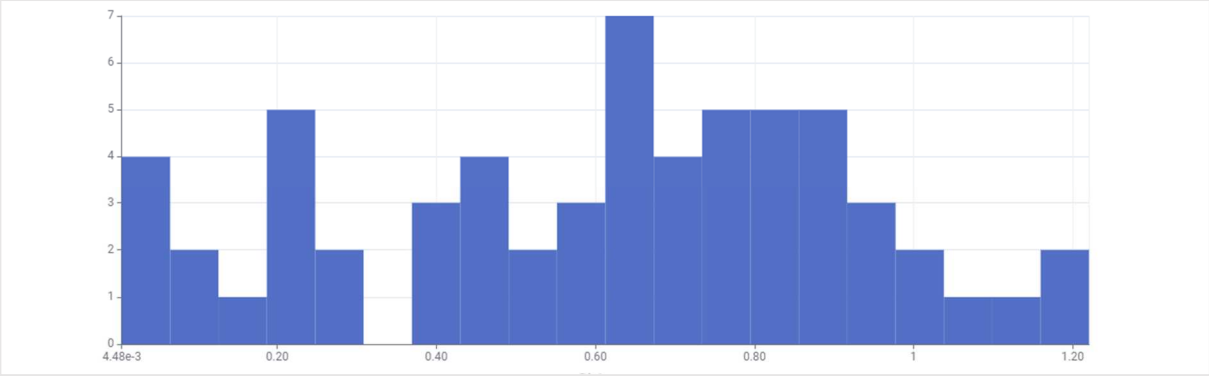



**Abbildung 26** Abstände der falsch negativen Beobachtungen zur Hyperebene

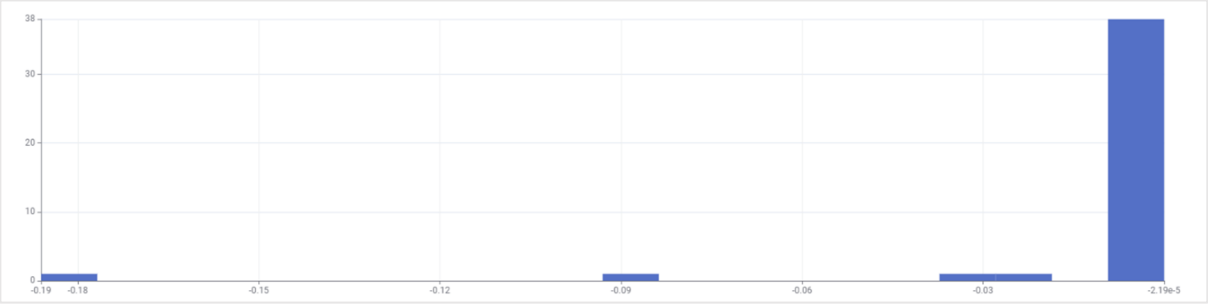



**Abbildung 27** Abstände der falsch positiven Beobachtungen zur Hyperebene

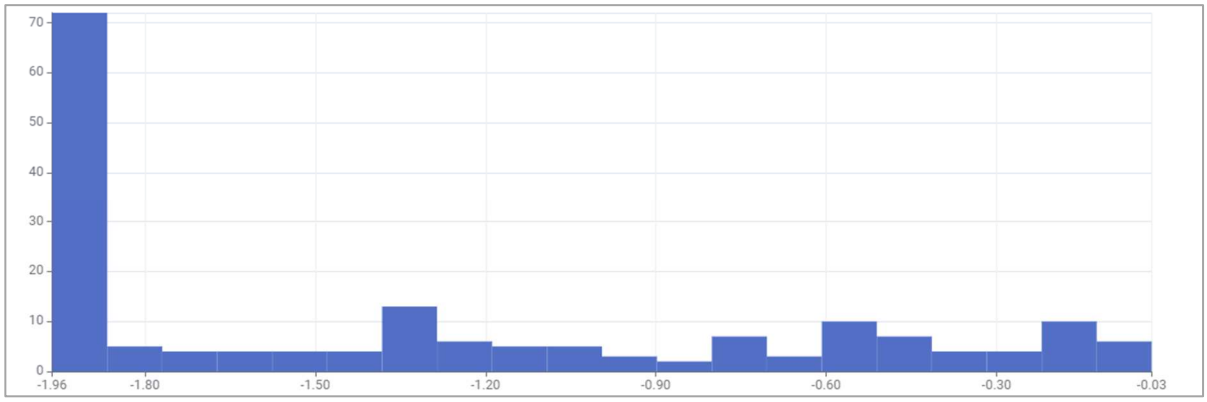
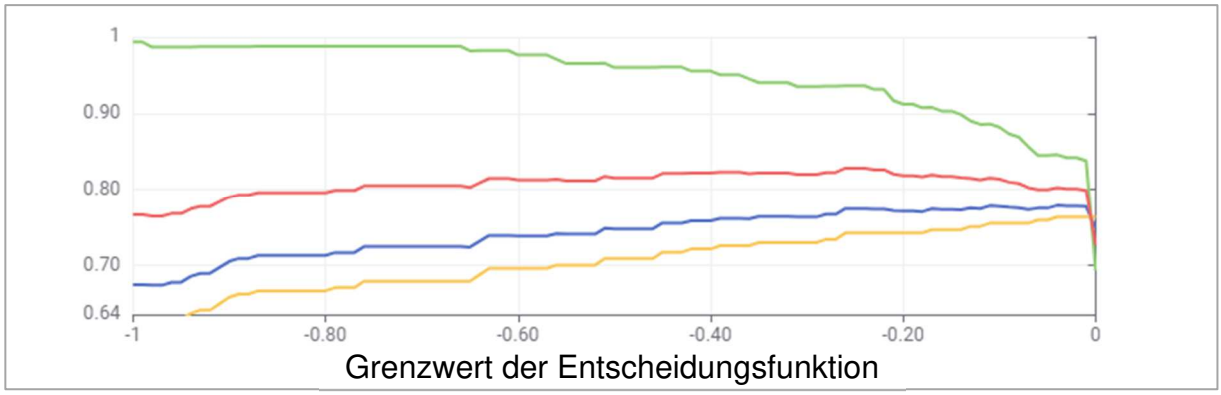



**Abbildung 28** Abstände der richtig positiven Beobachtungen zur Hyperebene

[111] Eigene Abbildungen

Die Idee ist es nun die Entscheidungsgrenze minimal zu verschieben, um eine bessere Modellperformance zu erreichen. Dafür wird der Abstand vom Modell ausgegeben und der Grenzwert gesucht, der den F2 Score maximiert. (vgl. Abbildung 29)




Sensitivität Präzision F1-Maß F2-Maß

**Abbildung 29** Anpassung des Grenzwertes der OCSVM zur Performanceoptimierung[112]

Der beste ermittelte Grenzwert liegt hierbei bei -0.01, wodurch eine Verbesserung des F2-Maßes des gesamten Datensatzes von 0.759 auf 0.788 erreicht werden konnte. Betrachtet man die Konfusionsmatrix nach Anpassung des Grenzwertes (vgl. Tabelle 8), dann zeigt sich eine Verbesserung bezüglich der falsch positiven Beobachtungen. Von vorher 46 sinkt die Anzahl auf 4 falsch positive Beobachtungen. Weiterhin ändert sich die Zahl der falsch negativen klassifizierten Fenster durch diese Änderung nicht.

Die Präzision des Modells wird also von 0.795 auf 0.978 verbessert, wobei die Sensitivität von 0.751 sich nicht verändert.

| | | *Vom Modell prognostiziert* | |
|---|---|---|---|
| | | **Anomalie** | **Normal** |
| *Tatsächlicher Wert* | **Anomalie** | 178 | 59 |
| | **Normal** | 4 | 14582 |

**Tabelle 8** Konfusionsmatrix nach angepasstem Grenzwert

Schaut man sich die falsch negativ klassifizierten Beobachtungen an fällt vor allem eine Gruppe an beschrifteten Anomalien auf. Von den 59 falsch negativ klassifizierten Daten liegen 36 innerhalb eines Zeitbereiches von einer Stunde. Schaut man sich in diesem Bereich Merkmal 3 an, (vgl. Abbildung 30) dann fällt dieser durch seine Ausbrüche nach oben auf.

[112] Eigene Abbildung

Da jedoch eine Fensterlänge von 2 Minuten gewählt wurde, enthalten nicht alle der beschrifteten Fenster diesen Ausbruch. Die Fenster werden neu beschriftet. Nun sind 21 der 36 Fenster als Anomalie beschriftet.

Zudem fällt auf, dass alle Fenster mit dieser Anomalie in einem, Streudiagramm der maximalen gegenüber der minimalen Abweichung von Merkmal 3 deutlich als eine zusammenhängende Gruppe erkennbar sind. (vgl. Abbildung 31)

**Abbildung 30** Falsch negativ klassifizierte Daten (grün) und falsch beschriftete Daten (rot)[113]

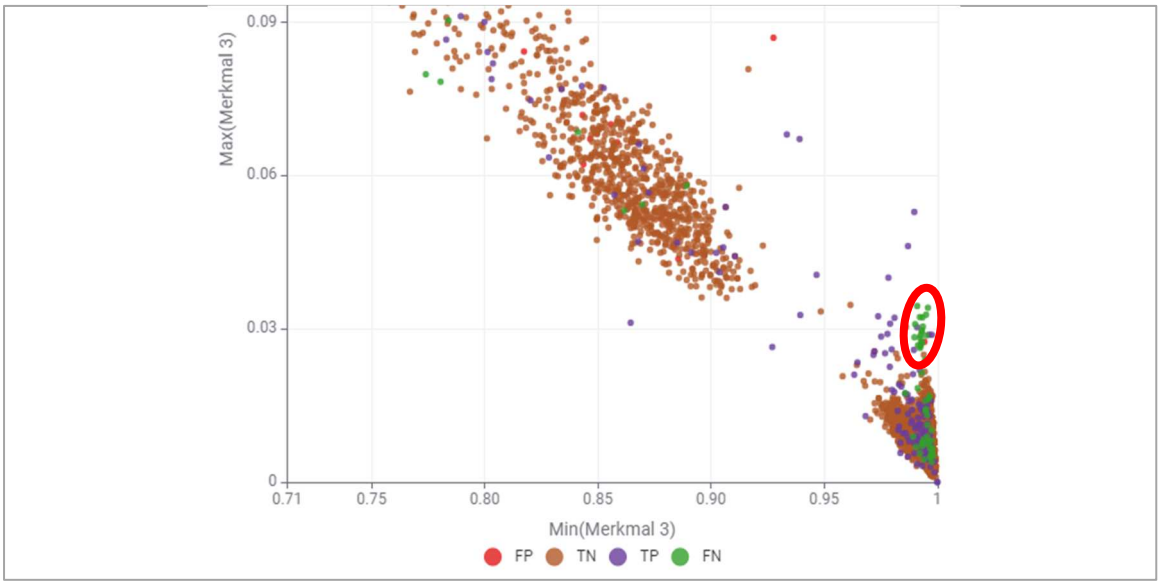



**Abbildung 31** Falsch negative Beobachtungen in Bezug auf Merkmal 3 [114]

[113] Eigene Abbildung
[114] Eigene Abbildung

Da diese jedoch sehr nahe an den Daten im Normalbereich liegen wurden diese nicht als Anomalie vom Modell prognostiziert. Die meisten dieser Beobachtungen haben einen großen positiven Abstand zur Trennebene. (vgl. Abbildung 32)

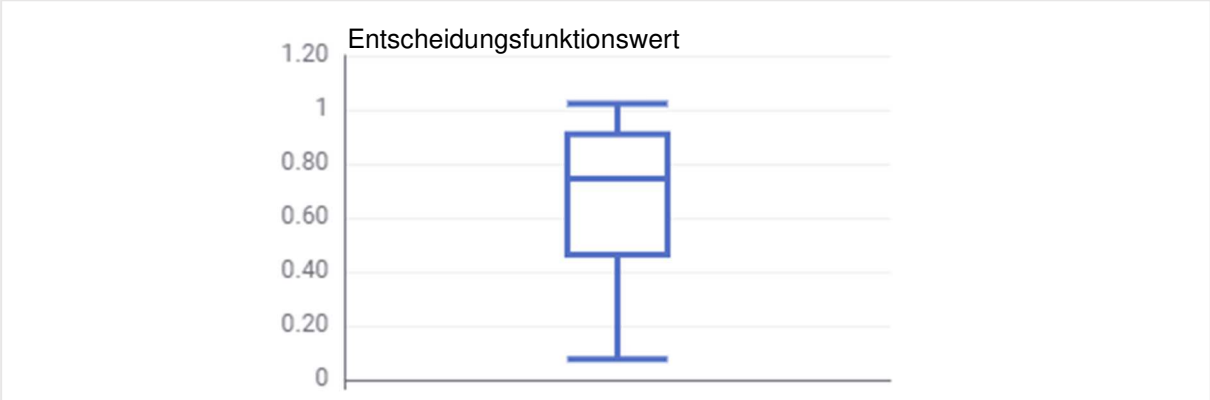



**Abbildung 32** Boxplot der Abstände zur Trennebene [115]

Zusätzlich nahe Datenpunkte zu dieser Gruppe vor dem Training zu entfernen führt hierbei nicht zu einer Performanceverbesserung. Jedoch könnte man getrennt von der AD solche Zustände über Entscheidungsgrenzen bzw. Klassifikationsalgorithmen bestimmen.

Das Modell wird deshalb nicht verändert. Jedoch werden die falsch beschrifteten Fenster, die diese Anomalie nicht enthalten neu gelabelt. Dadurch ergeben sich neue Performancemaße aus den Elementen der Konfusionsmatrix (vgl. Tabelle 9):

- Präzision: 0,978
- Sensitivität: 0,802
- F2-Maß: 0.832

| | | *Vom Modell prognostiziert* | |
|---|---|---|---|
| | | **Anomalie** | **Normal** |
| *Tatsächlicher Wert* | **Anomalie** | 178 | 44 |
| | **Normal** | 4 | 14597 |

**Tabelle 9** Konfusionsmatrix nach angepassten Labeln

In Bezug auf die beschrifteten Daten gibt es wenig Fehlalarme. Allerdings werden von den echten beschrifteten Anomalien 20% vom Modell verpasst. Dies kann einerseits an falschen Beschriftungen liegen. Andererseits kann es sein, dass das Modell durch die Nähe zu normalen Beobachtungen diese auch als normal bestimmt. Trotz der verpassten Anomalien ist die Modellleistung als gut einzuschätzen. Die hohe Präzision zeigt, dass das Modell verlässlich ist und wenig Fehlalarme erzeugt.

[115] Eigene Abbildung

# 5 Integration und Evaluation vom Anomalie-Erkennungs-system

## 5.1 Planung der Modellintegration in die produktive Umgebung

Nach dem Training muss das Modells in Produktionslinie integriert werden, um den Produktionsprozess zu überwachen. Einerseits können die Prozessdaten dafür durch zyklisches Polling oder Interrupt-basierte Mechanismen abgefragt werden.

Bei der Polling Variante werden neue Daten in festen Intervallen aus dem Datenbanksystem abgefragt.

Interrupt-basierte beziehungsweise ereignisgesteuerte Methoden basieren auf Unterbrechungen. Sobald neue Daten vorliegen, sendet beispielsweise ein Trigger der Datenbank ein Signal. Der aktuelle Prozess des Controllers wird unterbrochen und die Daten sofort an das Modell weitergegeben.

| METHODE | VORTEILE | NACHTEILE |
|---|---|---|
| **POLLINGKONZEPT** | • Einfach Implementierbar<br>• Geringe Infrastruktur<br>• Fehlerresistenter | • Verzögerte AD<br>• Ineffiziente Ressourcennutzung<br>• langsamer |
| **INTERRUPTKONZEPT** | • Schneller<br>• Priorisierungen Möglich<br>• Weniger Overhead | • Komplexe Umsetzung |

**Tabelle 10** Vergleich Polling und Interrupt[116]

Interrupts sind vor allem für zeitkritische Anwendung die bessere Wahl. Da es jedoch in diesem Fall genügt, dass die Daten in einem festen Intervall analysiert werden, wurde sich aufgrund der einfachen Umsetzung für die zyklische Datenabfrage entschieden.

Die Daten werden aus dem Datenbanksystem abgefragt und wie in Kapitel 4.3 dargestellt vorverarbeitet. Anschließend werden die Merkmalsvektoren der Fenster dem Modell übergeben und dieses klassifiziert die Beobachtung. (vgl. Abbildung 33)

[116] Vgl. LEHNGUTH TBPI Kapitel 2.4

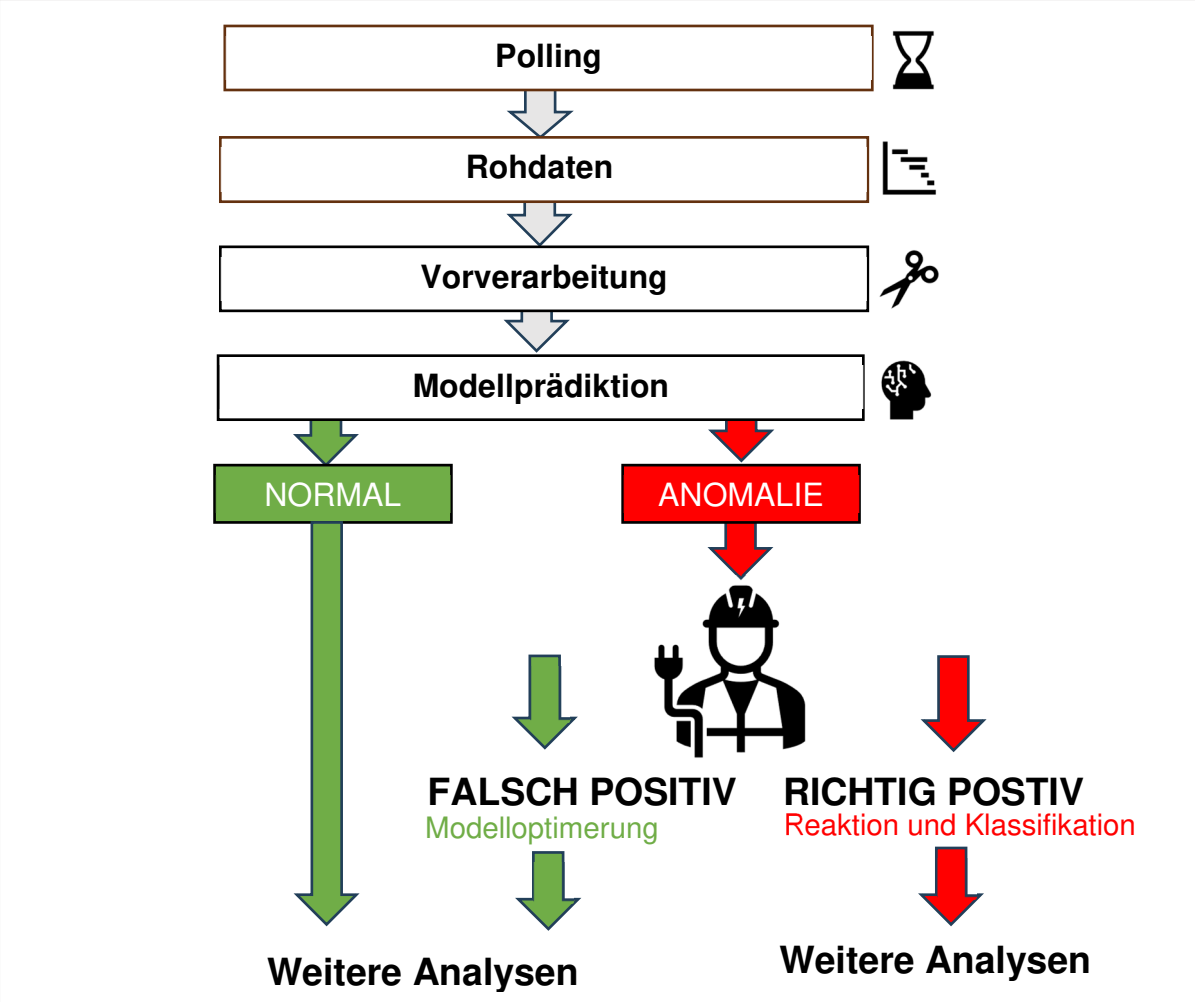



**Abbildung 33** Prozess der Anomaliedetektion[117]

Die Polling -Zykluszeit wird auf zehn Minuten festgelegt. Pro Zyklus werden die Daten aus dem letzten Stundeintervall aus der zugehörigen Datenrelation abgefragt. Dadurch entsteht eine große Menge an Redundanz. Es muss jedoch für die Ermittlung der Fenster mindestens einmal eine Bypassschaltung stattgefunden haben. Bei der Abfrage im Stundenintervall ist dies mindestens einmal der Fall.

Danach werden die Zeitreihen in die diskreten Fenster unterteilt, aus denen jeweils ein Merkmalsvektor bestimmt wird.

Anschließend erfolgt die Klassifikation von allen Fenstern, von denen noch keine Prädiktion vorliegt. Das Resultat des Modells wird in einer separaten Relation gespeichert. Diese umfasst Informationen über den Zeitbereich des Fensters, die Maschinenkennzeichnung und die bestimmte Klasse.

Zur kontinuierlichen Verbesserung des Modells, Charakterisierung der Anomalien und zur Ableitung operativer Handlungsempfehlungen ist eine Rückmeldung durch

[117] Eigene Abbildung

Prozesskundiges Personal essenziell. Zu diesem Zweck soll ein System geschaffen werden, dem Fachpersonal Anomalien mitzuteilen und Rückmeldungen zu den identifizierten Anomalien zu erhalten.

Um Handlungsempfehlungen aus den gefundenen Anomalien ableiten zu können, das Modell weiter zu optimieren und die Art und Schwere der Anomalien zu bestimmen zu können, sollen gefundene Anomalien von einem Prozesskundigen quittiert werden. Dabei soll vor allem dokumentiert werden, ob es sich bei einer Anomalie wirklich um eine Abweichung vom Normzustand handelt.

## 5.2 Verfügbarkeit und Nutzung des Modells

Zur Nutzbarmachung der Modellergebnisse erfolgt deren Visualisierung in einem interaktiven Dashboard. Dabei werden die originalen Zeitreihen der Sensordaten dargestellt und Anomalien visuell hervorgehoben. Zeitfenster mit einer Anomalie werden mit einem rot hinterlegten Bereich kenntlich gemacht.

Für diese Darstellung ist eine geeignete Verknüpfung von den Relationen der Sensordaten und AD erforderlich. Um zusammenhängende Anomalien visuell als Einheit darzustellen, werden aufeinanderfolgende Anomalien in einem durchgehenden Bereich zusammengeführt.

Dafür wird der erste Datenpunkt bei einem Wechsel von normalen in einen anormalen Zustand mit „1“ markiert. Erfolgt der Wechsel zurück in den Normalbereich, wird der letzte Datenpunkte mit „0“ gekennzeichnet. Durch rechteckige Interpolation nach vorne und füllen der Fläche unter der Kurve mit einem rot transparenten Hintergrund entsteht eine Visualisierung mit denen Anomalien rot markiert sind. (vgl. Abbildung 34)

**Abbildung 34** Markierte Anomalie im Dashboard[118]

Zusätzlich wird im Zeitreihendiagramm angezeigt, welches Werkstück, bei dem Auftreten einer Anomalie, beschichtet wurde. Dies ermöglicht es potenzielle Zusammenhänge zwischen aufgetretenen Anomalien und möglicher Qualitätsabweichungen einzelner Werkstücke zu identifizieren. Weiterhin können bei Bedarf weiterführende Untersuchungen durchgeführt werden, beispielsweise durch Laboranalysen.

Darüber hinaus können über das Dashboard auch Bilddaten der Bauteiloberflächen sowie aus der Bildverarbeitung ermittelten Merkmale angezeigt werden. Diese Unterstützen das Fachpersonal bei der Einschätzung möglicher Auswirkungen.

Zur kontinuierlichen Verbesserung des Modells und zur Handlungsempfehlungsableitung wird ein Rückmeldesystem für Werker implementiert. Dieses erlaubt die Validierung und Kommentierung erkannter Anomalien. Ziel ist es durch Expertenfeedback folgende Fragestellungen zu adressieren:

- Handelt es sich tatsächlich um eine relevante Anomalie?
- Gab es vergleichbare Anomalien in der Vergangenheit?
- Welche Ursachen könnten der Auffälligkeit zugrunde liegen?
- Wurde die Beschichtungsqualität beeinflusst?
- Gab es Auswirkungen auf die Betriebsmittel?

[118] Eigene Abbildung

- Wurden Maßnahmen ergriffen?

So ein Feedbacksystem bildet die Grundlage für weitere Untersuchungen, ein lernfähiges Anomalie-Management und zur Handlungsempfehlungsableitung.

## 5.3 Validierung und Kommentierung von erkannten Anomalien

Zur Bewertung der Modellgüte ist eine manuelle Rückmeldung zu den detektierten Anomalien angesehen. Wird eine Anomalie erkannt, so wird diese über das Dashboard kommuniziert.

Im Anschluss besteht die Möglichkeit die jeweilige Anomalie zu klassifizieren. Dabei stehen folgende Optionen zur Verfügung:

- Richtig positiv: Sollte erkannt werden
- Falsch positiv: Sollte nicht erkannt werden
- Nicht bewertet: Unsicher / Noch nicht untersucht

Zusätzlich kann die Fachkraft die detektierte Anomalie in einem Textfeld beschreiben. So sollen Informationen zu möglichen Ursachen, Kontextinformationen und weiteren Beobachtungen gesammelt werden. Ein weiteres Feld zur Eingabe von eventuell vorgenommenen Handlungsempfehlungen wird integriert. Hier können beispielsweise Wartungen oder Prüfentscheidungen hinterlegt werden. (vgl. Abbildung 35)

Wird eine Anomalie validiert und kommentiert, dann wird sie danach aus der Liste der offenen Anomalien entfernt.

**Abbildung 35** Kommentier- und Quittierung von klassifizierten Anomalien [119]

Durch dieses Vorgehen wird einerseits die Modellperformance am laufenden Prozess bewertet und andererseits die Grundlage für die kontinuierliche Weiterentwicklung des Erkennungssystems gebildet.

[119] Eigene Abbildung

## 5.4 Tests und Limitationen des entwickelten Systems

Zur Sicherstellung der Funktionsfähigkeit und Einsetzbarkeit werden verschiedene Tests durchgeführt. Diese lassen sich in Funktions- und Performancetests einzelner Komponenten (Komponententest) und des gesamten Systems (Integrationstest) unterteilen. Weiterhin wird ein Abnahmetest durchgeführt.

Folgende Funktionseinheiten werden unabhängig voneinander auf ihre Funktion geprüft und daraufhin auftretende Fehler behoben:

- Modell der OCSVM
- Datenvorverarbeitung
- Speicherung im Datenbanksystem
- Dashboard

Im Rahmen der Performancetests werden der Speicherplatz und die Ausführungszeit des Systems überprüft. Dies soll sicherstellen, dass die zur Verfügung stehenden Systemressourcen ausreichend sind, um alle eingehenden Daten innerhalb der Zykluszeit zuverlässig zu verarbeiten.

Die Performancetests werden direkt auf dem Zielsystem durchgeführt, also auf dem Rechner, auf dem die AD in die Produktionsumgebung integriert wird. Dadurch lassen sich praxisnahe Aussagen zur Stabilität und Skalierbarkeit treffen.

Die durchschnittliche Prädiktionszeit pro Beobachtung durch das trainierte Modell liegt bei 0,0003236 Sekunden. Der Speicherbedarf des Modells beträgt lediglich rund 16KB. Die Gesamtdauer eines vollständigen Zyklus, bestehend aus Datenabfrage, Vorverarbeitung, Modellprädiktion und Speicherung beträgt im Mittel 42 Sekunden. Die höchste Ausführungszeit betrug bei den Tests 49s. Damit bleibt die Bearbeitungszeit deutlich unter der Zykluszeit von 10 Minuten. Dies entspricht etwa einen Faktor von 1:14. Solange alle angebundenen Rahmensysteme (z.B. Datenbank, Netzwerk und Webserver) zuverlässig funktionieren, sollte somit eine vollständige Verarbeitung der Daten garantiert sein.

Im Rahmen der praxisnahen Bewertung wurde das System dem zuständigen Produktionspersonal vorgestellt. Dabei wurden folgende Rückmeldungen eingeholt.

Die Visualisierung erkannter Anomalien wurde als verständlich und hilfreich empfunden.

Die im Produktionsablauf erkannten Anomalien sind zum Großteil richtig erkannt worden. Probleme sehen die Werker noch bei der Klassifikation der Einschwingverhalten. Bei diesen kommt es öfters zu falsch positiven Klassifizierungen.

Die Möglichkeit zur Quittierung und Kommentierung wurde von den Werkern positiv bewertet. Hierdurch kann das System weiterentwickelt und Falschklassifikationen weiter reduziert werden.

Trotz der positiven Tests weist das System einige Limitationen auf.

- Das Modell der OCSVM wurde nur auf einer Maschine trainiert. Durch unterschiedliche Betriebsmitteleigenschaften kommt es auf anderen Maschinen häufiger zu Falschklassifikationen.
- Die OCSVM bietet lediglich eine binäre Entscheidung. Handlungsempfehlungen können noch nicht abgeleitet werden. Mit diesem System wurde jedoch die Grundlage gelegt.
- Das Modell kann sich nicht selbst anpassen und muss manuell nachtrainiert werden.
- Das System wurde nur für einen spezifischen Teilbereich (Pulverinjektion) des APS entwickelt. Die Übertragbarkeit auf andere Teilsysteme bleibt offen.

Die absolvierten Tests haben gezeigt, dass das System funktionsfähig und praxisnah einsetzbar ist. Gleichzeitig wurden mehrere Limitationen identifiziert. Diese Erkenntnisse bilden die Grundlage für eine gezielte Weiterentwicklung.

Im Folgenden Kapitel werden die wichtigsten Ergebnisse und Erkenntnisse dieser Arbeit dargestellt und kritisch reflektiert. Weiterhin wird in Ausblick gegeben, wie mögliche Weiterentwicklungen aussehen könnten.

## 5.5 Typische Anomaliebeispiele

In der nachfolgenden Abbildung 36 sind exemplarisch erkannte Anomalien dargestellt, wobei diese farblich Hervorgehoben sind. Diese sollen die unterschiedliche Erscheinungsform der erkannten anomalen Muster verdeutlichen. Die dargestellten Fälle wurden vom System automatisch erkannt.

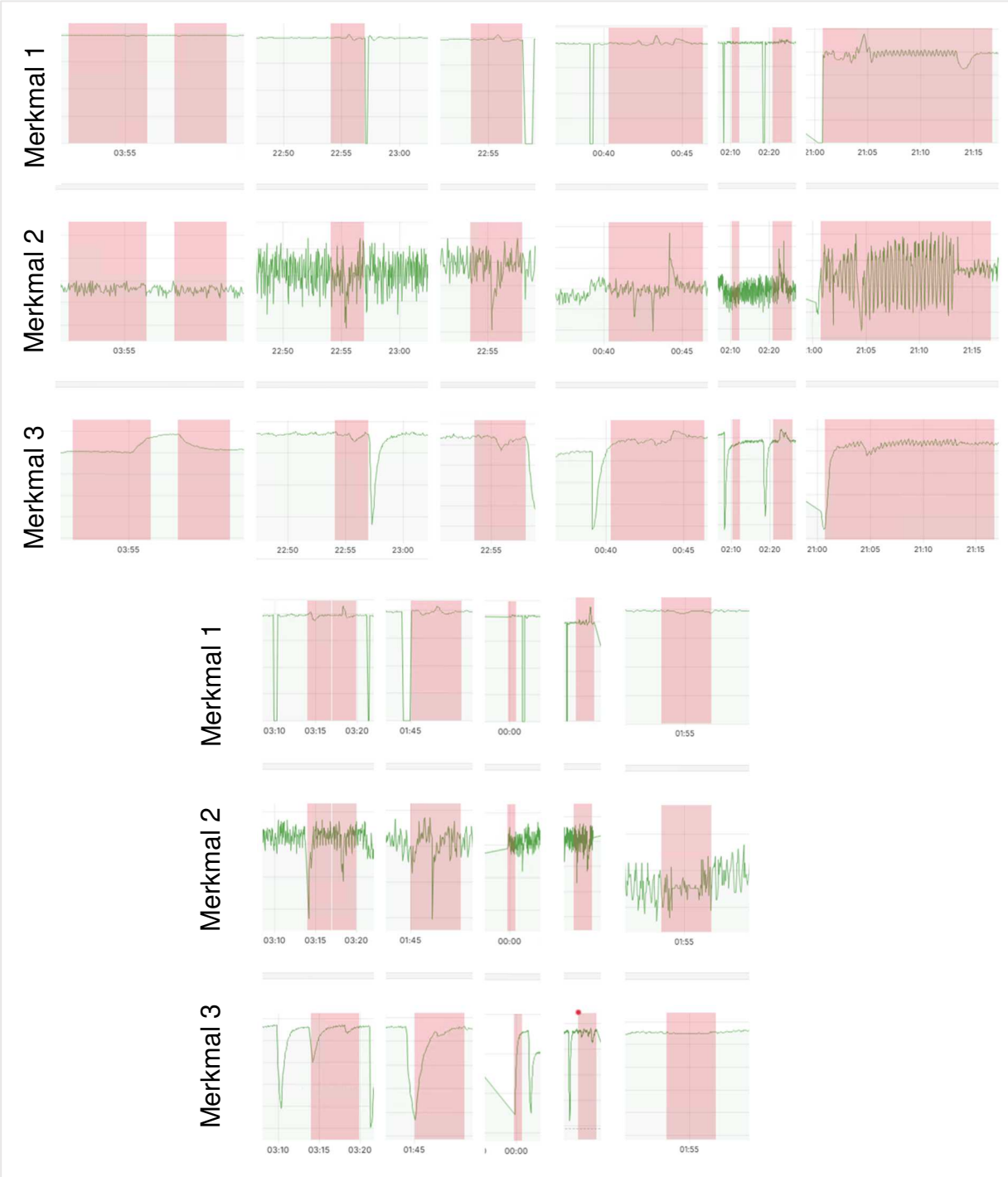



**Abbildung 36** Vom AD-System erkannte Anomalien

# 6 Fazit und Ausblick

## 6.1 Zusammenfassung der wichtigsten Ergebnisse

Im Rahmen dieser Arbeit wurde ein datengetriebenes AD-System für die Überwachung der Pulverinjektion im APS-Prozess zur Innenbeschichtung von ZKG's entwickelt, implementiert und evaluiert. Ziel war es, den Bogen zwischen den theoretischen Grundlagen der AD und der industriell einsatzfähigen Anwendung zu spannen.

Auf Grundlage eines methodischen Überblickes über verschiedene Verfahren wurde die OCSVM als zentrales Modell zur AD aus dem Bereich des ML ausgewählt und mathematisch beschrieben.

Durch die Datenvorverarbeitung und Beschriftung des Datensatzes wurde die Grundlage für die Modellierung der AD geschaffen. Die erzeugten Label ermöglichen eine gezielte Optimierung von AD-Modellen und bilden zudem die Basis für den Vergleich verschiedener Methoden.

Die OCSVM zeigt eine hohe Güte und vor allem eine sehr hohe Präzision auf. Damit hat sich gezeigt, dass die OCSVM als Modell für den prototypischen und praktischen Einsatz eignet. Weiterhin konnte damit bestätigt werden, dass sich ein datengetriebenes AD-System erfolgreich in der reellen Produktionsumgebung des APS integrieren lässt.

Ein wesentliches Ergebnis war zudem die Darstellung der Modellergebnisse in einem dynamischen Dashboard. Neben der farblichen Hervorhebung erkannter Anomalien wurde ein Quittier- und Kommentiersystem implementiert. Dies soll einerseits sicherstellen, dass Anomalien vom Werker war genommen werden und andererseits dem Fachpersonal die Möglichkeit geben, erkannte Anomalien zu validieren und bei Bedarf zu kommentieren. Dieses Feedback bildet die Grundlage für die kontinuierliche Verbesserung des Modells und Unterstützen die geplante Handlungsempfehlungsableitung.

Zusammenfassend lässt sich sagen, dass das entwickelte System soliden Ansatz darstellt, um Anomalien im Teilsystem der Pulverinjektion zu erkennen. Jedoch gibt es noch Aspekte, welche zukünftig weiter optimiert werden können. Die Nutzung und Optimierung des Modells hängt jedoch von seiner regelmäßigen Anpassung und der Einbindung von Fachpersonal ab.

## 6.2 Kritische Reflexion und praxisnahe Bewertung

Die in dieser Arbeit entwickelte Lösung, zeigt wie ein datengetriebenes AD-System in eine reelle Produktionsumgebung integriert werden kann. Die durchgeführten Tests haben die Praxistauglichkeit, Funktionalität und Genauigkeit des entwickelten Systems belegt. Dennoch lassen sich aus dem Ergebnis mehrere kritische Punkte ableiten.

Die OCSVM wurde bewusst als methodischer Ansatz ausgewählt und zeigte in Bezug auf den beschrifteten Datensatz eine hohe Modellgüte. Allerdings wurden andere potenziell geeignete Verfahren aus dem Bereich des ML und DL im Rahmen dieser Arbeit bisher nicht untersucht. Ein Vergleich dieser Methoden könnte wertvolle Erkenntnisse liefern.

Die Datenbeschriftung erfolgte auf Grundlage eines vorverarbeiteten Datensatzes mit vordefinierter Schrittweite und Fenstergröße. Änderung dieser Parameter würden den Datensatz für die Evaluierung der Modellperformance unbrauchbar machen, da die Annotation auf Basis der Fenster und nicht punktbasiert vorgenommen wurde. Zudem sind die Datenpunkte eines Fensters auf den Mittelwert des jeweiligen Fensters zentriert. Dadurch erkennt das aktuelle System keine langfristigen Trendbewegungen, welche einen entscheidenden Einfluss auf die Modellqualität haben können

Die ersten Tests in der realen Produktionsumgebung haben gezeigt, dass Anomalien zuverlässig im Dashboard dargestellt und von den Werkern wahrgenommen werden. Die Möglichkeit zur Kommentierung und Validierung bietet die Grundlage für die weitere Fehlererkennung und Handlungsempfehlungsableitung. Die Effektivität dieses Systems hängt allerdings stark von der aktiven Nutzung durch das Werkspersonal ab. Eine geringe Beteiligung kann die kontinuierliche Verbesserung hierbei erheblich einschränken.

Darüber hinaus werden aufeinanderfolgende anormale Fenster im Dashboard zwar als zusammenhängende Einheit dargestellt, müssen jedoch derzeit individuell bewertet werden. Für eine effizientere Bedienung wäre zukünftig ein Mechanismus wünschenswert die zusammenhängenden Segmente gemeinsam validieren lässt.

Insgesamt belegt die Umsetzung, dass das System technisch zuverlässig arbeitet und eine wertvolle Unterstützung der Prozessüberwachung bietet. Für eine langfristige Nutzung sind jedoch noch Erweiterungen in der Methodenvielfalt, der Trend- und Kontextanalysen sowie Interaktionsgestaltungen erforderlich.

## 6.3 Weiterentwicklungsmöglichkeiten und Forschungsperspektiven

Aus den Ergebnissen dieser Arbeit ergeben sich mehrere aufeinander aufbauende Entwicklungsmöglichkeiten, das System robuster, aussagekräftiger und nutzungsnäher zu machen.

Kurzfristig sollte ein strukturierter Verfahrensvergleich erfolgen. Ein Vergleich der OCSVM mit anderen Verfahren der AD wie z.B. Isolationswälder oder Autoencodern könnte die Modellperformance weiter verbessern. Auch hybride Ansätze wären denkbar, um die Stärken verschiedener Modelle zu kombinieren.

Ziel ist nicht nur eine höhere Trefferquote, sondern auch eine stabilere Falschalarmraten unter wechselnden Bedingungen. Hier ist vor allem die Anpassung des Modells auf mehrere Systeme desselben Beschichtungsprozesses notwendig.

Da relevante Abweichungen häufig schleichend entstehen, sollten in weiteren Analysen Trendeffekte explizit berücksichtigt werden. Mittels Drift- und Änderungsindikatoren könnten frühzeitig Warnungen generiert werden. Da sich solche Trendbewegungen meist auf die Bauteilqualität auswirken sollte auch ein Zusammenhang zwischen Trendbewegungen in den Qualitätsmerkmalen und den Prozessparametern untersucht werden.

Erkannte Anomalien könnten mit Clusteringverfahren automatisch gruppiert werden. Wiederkehrende Muster erhalten damit genauere Label (z.B. Düse verstopft o. Pulverfluss schwankt). Dies würde den Kommentierungsprozess effizienter gestalten. Die vom Produktionspersonal kommentierten Anomalien können hierfür als Grundlage genutzt werden. Aus verschiedenen Typen an Anomalien könnten so schneller Vorschläge, Ursachen oder Regeln abgeleitet werden.

Das Kommentiersystem könnte zudem zielgerechter genutzt werden. Auf Basis einer Wahrscheinlichkeit für bestimmte Fehlerbilder könnten gezielt unsichere Fälle abgefragt werden. Um den geschäftlichen Mehrwehrt zu belegen, sollten Anomalien mit Qualitätsmerkmalen verknüpft werden. Es muss hierbei untersucht werden, welche Anomalien qualitätsrelevant sind.

Auf Basis der daraus gewonnen Kenntnissen könnten bestimmte Typen zu verschiedenen Alarmen und Warnungen führen, welche durch autorisiertes Personal wahrgenommen werden müssen. Die Quittierung im bisherigen System ist aufgrund des prototypischen Einsatzes nicht verpflichtend.

Ein denkbarer Schritt wäre ein Transfer des Systems auf andere Teilsysteme. So könnte die Generalisierbarkeit nachgewiesen werden. Denkbar wäre zum Beispiel die Analyse der Leistungsparameter des Beschichtungsprozess.

Mit diesen Schritten entwickelt sich der aktuelle Prototyp zu einem System, das nicht nur Anomalien erkennt, sondern deren Bedeutung für die Qualität sichtbar macht und mittelfristig automatisch konkrete Handlungsempfehlungen ableiten kann. Insgesamt zeigt sich, dass das entwickelte System, das Potenzial hat, die Grundlage für eine vorrausschauende und qualitätsorientierte Prozessüberwachung zu bilden.

# Quellenverzeichnis

AKSHAY B.; U. A.: Machine learning – A comprehensive beginner's guide. First edition. Boca Raton, 2025

ALPAYDIN, Ethem: Maschinelles Lernen: De Gruyter Studium. 3., aktualisierte und erweiterte Auflage. Berlin, 2022

BELAY, Mohammed A.; u. a. : Unsupervised Anomaly Detection for IoT-Based Multivariate Time Series: Existing Solutions, Performance Analysis and Future Directions. In: Sensors (Basel, Switzerland), Jg. 23, 2023, H. 5, S. 1 - 24. DOI: 10.3390/s23052844

BERRAR, Daniel: Cross-validation, 2019

BERWICK, R.: An Idiot’s guide to Support vector machines (SVMs). In: https://web.mit.edu/6.034/wwwbob/svm-notes-long-08.pdf (17.07.2025)

BISHOP, Christopher M.: Pattern Recognition and Machine Learning, 2006

CHANDOLA, Varun; BANERJEE, Arindam; KUMAR, Vipin: Anomaly detection: A survey. In: ACM Computing Surveys, Jg. 41, 2009, H. 3, S. 1 - 58. DOI: 10.1145/1541880.1541882

FOORTHUIS, R.: On the Nature and Types of Anomalies: A Review of Deviations in Data. In: International Journal of Data Science and Analytics, 2021

FRANC, Vojtěch; HLAVÁČ, Václav: An iterative algorithm learning the maximal margin classifier. In: Pattern Recognition, Jg. 36, 2003, H. 9, S. 1985 - 1996. DOI: 10.1016/S0031-3203(03)00060-8

Geschichtete bzw. stratifizierte Stichprobe | Lehrbuch Psychologie, 2025. In: https://www.lehrbuch-psychologie.springernature.com/glossar/geschichtete-bzw-stratifizierte-stichprobe (06.08.2025)

GOODFELLOW, Ian; COURVILLE, Aaron; BENGIO, Yoshua: Deep learning: Adaptive computation and machine learning. Cambridge, Massachusetts, 2016

GOULEBE, rashid: Erkennung von Anomalien in Zeitreihen durch One Class SVM und K-Means Clustering FH Aachen. Aachen, 2024

HEINE, Moreen; u. a.: Künstliche Intelligenz in öffentlichen Verwaltungen – Grundlagen, Chancen, Herausforderungen und Einsatzszenarien: Edition eGov-Campus. Wiesbaden, 2023

HEMPEL, A.-J.: Vorlesung MSR - Teilbereich "Regelungstechnik" DHSN Glauchau

HEMPEL, A.-J.: Vorlesung TBPI - Teilbereich "Identifikation" DHSN Glauchau

HOCHSCHULE TRIER: Grundlagen Neuronaler Netze. In: https://www.hochschule-trier.de/fileadmin/Hauptcampus/Fachbereich_Informatik/Fernstudium/Dokumente/Leseproben/bdl_grundlagen.pdf

HODGE, Victoria J.; AUSTIN, Jim: A Survey of Outlier Detection Methodologies. In: Artificial Intelligence Review, Jg. 22, 2004, H. 2, S. 85 - 126. DOI: 10.1007/s10462-004-4304-y

HU, Jie; u. a. : Prediction Performance Improvement via Anomaly Detection and Correction of Actual Production Data in Iron Ore Sintering Process. In: IEEE Transactions on Industrial Informatics, Jg. 16, 2020, H. 12, S. 7602 - 7612. DOI: 10.1109/TII.2020.2979465

HUDE, Marlis von der: Predictive analytics und data mining – Eine Einführung mit R: Lehrbuch. Wiesbaden, 2020

HURST, Marco; WENTZIEN, Marcel; SCHMALZRIED, Dirk: Erklärbare künstliche Intelligenz im CRISP-DM-Prozess. In: Wirtschaftsinformatik & Management, Jg. 16, 2024, H. 5-6, S. 250 - 259. DOI: 10.1365/s35764-024-00533-2

INOUE, Jun; u. a., (2017): Anomaly Detection for a Water Treatment System Using Unsupervised Machine Learning. In: GOTTUMUKKALA, Raju [Hrsg.]: 17th IEEE International Conference on Data Mining workshops – 18-21 November 2017, New Orleans, Louisiana : proceedings. 2017 IEEE International Conference on Data Mining Workshops (ICDMW). New Orleans, LA, 11/18/2017 - 11/21/2017. Institute of Electrical and Electronics Engineers; IEEE Computer Society. Piscataway, NJ: IEEE, S. 1058 - 1065

ISERMANN, R.: Fahrdynamik-Regelung: Modellbildung, Fahrerassistenzsysteme, Mechatronik., 2006

JAYASWAL, Vaibhav: Performance Metrics: Confusion matrix, Precision, Recall, and F1 Score | Towards Data Science, 2020. In: https://towardsdatascience.com/performance-metrics-confusion-matrix-precision-recall-and-f1-score-a8fe076a2262/ (24.07.2025)

KAMPS, U.: Definition: Normalverteilung. In: Springer Fachmedien Wiesbaden GmbH, 16.02.2018. In: https://wirtschaftslexikon.gabler.de/definition/normalverteilung-39769/version-133790

KASTE, Jonas: Künstliche Neuronale Netzwerke Zur Adaptiven Fahrdynamikregelung. v.171: AutoUni - Schriftenreihe Series. 1st ed. Wiesbaden, 2024

KUNZE, Katja: Support Vector Machines, Kernels, 2004. In: https://campar.in.tum.de/twiki/pub/Far/MachineLearningWiSe2003/kunze_ausarbeitung.pdf (18.07.2025)

LAURIKKALA, J.; JUHOLA, M.; KENTALA, E.: Informal identification of outliers in medical data, 2000, S. 20 - 24

LEE, Namgil; YANG, Heejung; YOO, Hojin, (2021): A surrogate loss function for optimization of $F_β$ score in binary classification with imbalanced data. In: http://arxiv.org/pdf/2104.01459

LEHNGUTH, T.: Vorlesung TBPI - Teilbereich "Prozessinformatik" DHSN Glauchau

LIASHCHYNSKYI, Petro; LIASHCHYNSKYI, Pavlo, (2019): Grid Search, Random Search, Genetic Algorithm: A Big Comparison for NAS. In: http://arxiv.org/pdf/1912.06059

LOHNINGER, H.: Ausreißertests - Grundregeln, 2021. In: http://www.statistics4u.info/fundstat_germ/cc_outlier_tests_4sigma.html (11.07.2025)

MUNIR, Mohsin; u. a. : DeepAnT: A Deep Learning Approach for Unsupervised Anomaly Detection in Time Series. In: IEEE Access, Jg. 7, 2019, S. 1991 - 2005. DOI: 10.1109/ACCESS.2018.2886457

NIELSEN, Michael A.: Neural Networks and Deep Learning, 2015

OERLIKON GROUP: Die 3 wichtigsten thermischen Spritzverfahren, die Sie kennen sollten, 2024. In: https://www.oerlikon.com/metco/de/loesungen-und-technologien/was-ist-thermisches-spritzen/top-3-thermische-spritzverfahren/ (04.08.2025)

PASCHEN, Friedrich: Über die zum Funkenübergang in Luft, Wasserstoff und Kohlensäure bei verschiedenen Drucken erforderliche Potentialdifferenz. In: Annalen der Physik, Bd. 273, 1889

POENISCH, J.: Vortrag SVM. TU Chemnitz. In: https://www.tu-chemnitz.de/urz/ittime/documents/Vortrag_Jens_Poenisch_SVM.pdf (17.07.2025)

RAFFERTY, Mark; u. a. : Real-Time Multiple Event Detection and Classification Using Moving Window PCA. In: IEEE Transactions on Smart Grid, Jg. 7, 2016, H. 5, S. 2537 - 2548. DOI: 10.1109/TSG.2016.2559444

REUß, Andreas; ZWIESLER, Hans-Joachim: Ein generisches Kreislaufmodell zur Einbettung von Data-Mining-Analysen in die Gesch¨aftsprozesse von Unternehmen – mit einem Fallbeispiel aus der Versicherung

RICHTER, Johannes: Entwicklung einer Prozessregelung für das atmosphärische Plasmaspritzen zur Kompensation elektrodenverschleißbedingter Effekte, 2014

RUFF, L.: A Unifying Review of Deep and Shallow Anomaly Detection. In: Proceedings of the IEEE, Mai 2021

RUNKLER, Thomas A.: Data mining – Modelle und Algorithmen intelligenter Datenanalyse: Lehrbuch. 2., aktualisierte Auflage. Wiesbaden, 2015

SHANG, Wenli; u. a. : Intrusion detection algorithm based on OCSVM in industrial control system. In: Security and Communication Networks, Jg. 9, 2016, H. 10, S. 1040 - 1049. DOI: 10.1002/sec.1398

STOETZER, Matthias-W.: Zeitreihenanalyse und dynamische Modelle. In: STOETZER, Matthias-Wolfgang [Hrsg.]: Komplexe Verfahren. Berlin, 2020, S. 109 - 225

DIN EN ISO 14917, August 2017: Thermisches Spritzen – Begriffe, Einteilung

WANG, Zhongfeng; u. a. : Power System Anomaly Detection Based on OCSVM Optimized by Improved Particle Swarm Optimization. In: IEEE Access, Jg. 7, 2019, S. 181580 - 181588. DOI: 10.1109/ACCESS.2019.2959699

Was ist halbüberwachtes Lernen? | IBM, 2025. In: https://www.ibm.com/de-de/think/topics/semi-supervised-learning (05.08.2025)

WINTER, P.: Anomalieerkennung in Computernetzen. Datenschutz und Datensicherheit. In: DuD

WITTE, Kerstin: Angewandte Statistik in der Bewegungswissenschaft. Band 3: Lehrbuch. Berlin, 2019

ZONTA, Tiago; u. a. : Predictive maintenance in the Industry 4.0: A systematic literature review. In: Computers & Industrial Engineering, Jg. 150, 2020, S. 106889. DOI: 10.1016/j.cie.2020.106889

# Eidesstattliche Erklärung

Ich erkläre an Eides statt,

dass ich die vorliegende Arbeit selbständig und nur unter Verwendung der angegebenen Quellen und Hilfsmittel angefertigt habe.

Die aus fremden Quellen direkt oder indirekt übernommenen Stellen sind als solche kenntlich gemacht. Bestandteile der Arbeit, die mittels Künstlicher Intelligenz entstanden sind, wurden ausdrücklich gekennzeichnet.

Die Zustimmung des/der beteiligten Unternehmen/s zur Verwendung betrieblicher Unterlagen habe ich eingeholt.

Die Arbeit wurde bisher in gleicher oder ähnlicher Form weder veröffentlicht noch einer anderen Prüfungsbehörde/-stelle vorgelegt.

Schmidt, Etienne
…….
Name, Vorname Verfassender

Glauchau, 13.08.2025
…… ……..
Ort, Datum Abgabetermin Unterschrift Verfassender