

# Konzeption und Aufbau einer Trainingsinfrastruktur für virtuelle Sensoren

Malte Hoffmann

Technische Informatik

Hochschule  
Albstadt-Sigmaringen  
Gartenstr. 15  
72458 Albstadt

Derk Rembold

Technische Informatik

Hochschule  
Albstadt-Sigmaringen  
Gartenstr. 15  
72458 Albstadt  
E-Mail: [rembold@hs-albsig.de](mailto:rembold@hs-albsig.de)

Bernd Stauß

Wirtschaftsinformatik

Hochschule  
Albstadt-Sigmaringen  
Johannesstr. 6  
72458 Albstadt  
E-Mail: [staussb@hs-albsig.de](mailto:staussb@hs-albsig.de)

Stefan Schwarzkopf

Technische Informatik

42aaS GmbH  
Vor dem Kreuzweg 17  
7070 Tübingen  
E-Mail: [stefan.schwarzkopf@42aas.de](mailto:stefan.schwarzkopf@42aas.de)

Mario Jacobi

Technische Informatik

Hochschule  
Albstadt-Sigmaringen  
Gartenstr. 15  
72458 Albstadt

## ABSTRACT

Im Rahmen des Forschungsprojekts KI-VISOPRO (Entwicklung KI-basierter virtueller Sensoren zur Prozesssteuerung) wurde ein Sensorsystem entwickelt, welches komplexe Prozessgrößen bestimmen kann, die mit herkömmlichen Sensoren nicht trivial messbar sind. Hierzu wurde ein Ansatz der künstlichen Intelligenz gewählt, der dem supervised Deep Learning zuzuordnen ist. Ein essenzieller Bestandteil des Projektes ist der Aufbau und die Inbetriebnahme eines Demonstrators zum strukturierten Sammeln der Prozess- und Sensordaten mittels eines Enterprise Resource Planning (ERP) Systems. Die ermittelten Prozessgrößen können dabei auch wieder in den Prozess und das ERP-System zurückgespielt und dort zur Prozesssteuerung oder der Optimierung von Prozessschritten und -parametern verwendet werden.

## Motivation

Sollen physikalische Werte als Ziel- oder Leitgrößen in Prozessen verwendet werden, so müssen diese zuverlässig gemessen werden können. Dies erfordert oftmals teure und im Betrieb, Einbau und Wartung aufwendige Sensoren oder Messprozesse. Ein Beispiel hierfür ist der Messprozess zur Bestimmung der absoluten Feuchtigkeit in Baustoffproben oder Lebensmitteln wie Kaffeebohnen.

In beiden Fällen wird oftmals die Darr-Methode (gravimetrische Messung) eingesetzt. Dabei wird eine Probe gewogen, für eine bestimmte Zeit in eine Trockenkammer gestellt und letztendlich nochmals gewogen. Aus dem gemessenen Gewichtsverlust kann auf die anfangs enthaltene Feuchtigkeit geschlossen werden. Dieser aufwendige Messprozess benötigt einen manuellen Eingriff, welcher den Gesamt-Produktionsprozess verzögern kann.

Ein weiteres Beispiel findet man in der Bestimmung des Zuckergehalts während enzymatischer Umsetzungs-

bzw. Fermentationsprozesse, wie sie z.B. bei der Bierherstellung auftreten. Bei der Bestimmung des Zuckergehalts in Flüssigkeiten wird üblicherweise das Refraktometer oder vorzugsweise die Spindel (Saccharometer) verwendet. Die Spindel selbst zeigt dabei nicht direkt den Zuckergehalt an, sondern die Dichte der Flüssigkeit. Man weiß, dass die Dichte des reinen Wassers bei ca. 4°C 1000 kg/m<sup>3</sup> beträgt. Der Zucker in der Flüssigkeit beeinflusst die Dichte der Lösung, woraus sich von der Dichtemessung auf den Zuckergehalt schließen lässt. Auch hier ergibt sich ein manueller Eingriff in den Prozessablauf, da Proben entnommen werden müssen. Die Spindel-Messung erscheint auf den ersten Blick verhältnismäßig unproblematisch, allerdings muss die Probe vor der Messung auf eine bestimmte Temperatur (meist 20°C, selten 70°C) gekühlt bzw. geheizt werden, da Saccharometer auf diese Temperaturen kalibriert sind. Hierdurch entsteht eine weitere zeitliche Verschiebung, bevor der Messwert als Prozessparameter verwendet werden kann.

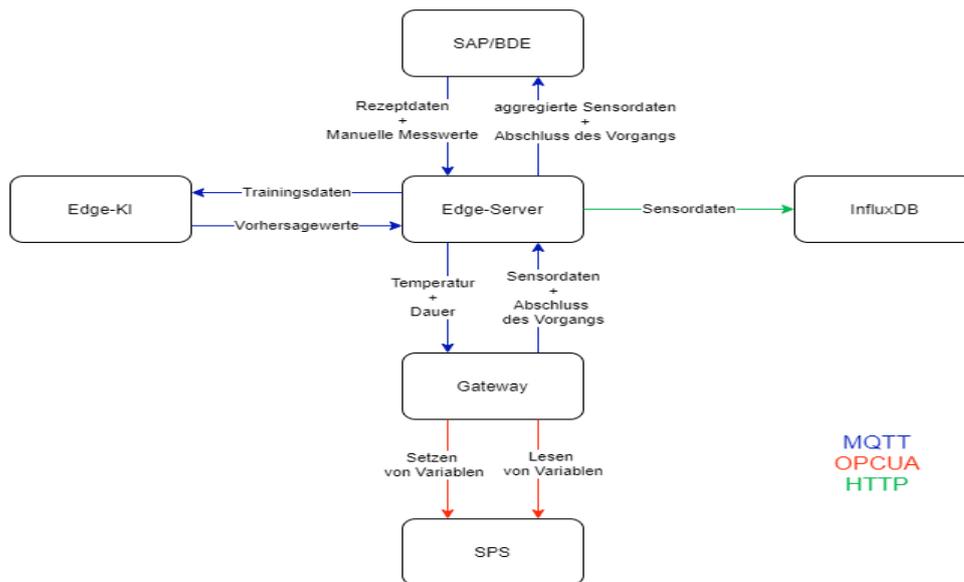


Abbildung 1: KI-VISOPRO vereinfachte Infrastruktur

Bei kleineren Temperaturunterschieden können Umrechnungstabellen helfen, jedoch erfordern diese ein sehr genaues Ablesen. Zudem können Störfaktoren wie Schwebstoffe oder Auftrieb durch Kohlensäure die Messung stark beeinflussen.

Für die Bestimmung des Zuckergehaltes gibt es bereits Geräte, die „inline“ Prüfungen, d.h. ohne explizite Probenahme, ermöglichen. Dabei wird unter anderem Ultraschall eingesetzt, um anhand von Laufzeitunterschieden und Schallabsorption auf die Dichte der Flüssigkeit zu schließen. Auch hier wird ähnlich wie bei der Spindel nicht direkt der Zuckergehalt, sondern die Dichte der Flüssigkeit gemessen. Diese Geräte sind jedoch sehr teuer und müssen regelmäßig von Fachpersonal gepflegt werden.

Im Rahmen des Projektes KI-VISOPRO wurden insbesondere die Messung des Zuckergehalts von Flüssigkeiten, am Beispiel Bierwürze, in den Mittelpunkt der Untersuchungen gerückt.

### Beschreibung des Gesamtsystems

Abbildung 1 zeigt auf unterster Ebene den technischen Prozess, welcher im konkreten Fall auf einer Bierbrauanlage der Firma Speidel durchgeführt wird. Der technische Prozess wird über eine speicherprogrammierbare Steuerung (SPS) gesteuert. Über die SPS können einige Sensorwerte, z.B. Druckwerte, Stromverbrauchswerte der Aktoren und Temperaturwerte, ausgelesen werden. Zudem wird ein Sensormodul der Firma ifm verwendet, um das Abnehmen des Deckels des Bierbraukessels zu erkennen. Letztendlich gibt es noch eine Waage der Firma Kern, um das enthaltene Gewicht des Bierbraukessels zu messen. Die verschiedenen Komponenten werden in einem Gateway über unterschiedliche Protokolle und Technologien, bspw. OPC-UA, RS232 oder REST-Interface, gesammelt und zentral versendet. Durch dieses

Gateway wird zusätzlich die SPS von der Außenwelt getrennt, so dass von außen über das Internet nicht direkt auf die SPS zugegriffen werden kann. Zur Kommunikation mit der Außenwelt wird ein MQTT-Broker verwendet, der die Nachrichten an seine Subscriber verteilt. Zu diesen Subscribern gehört der Edge-Server, welcher die Nachrichten verarbeitet, die Kommunikation der Komponenten überwacht und die Sensorwerte in die InfluxDB Datenbank abspeichert. Der Edge-Server enthält eine State-Machine, welche den aktuellen Zustand eines aktiven Brauprozesses abbildet. Über die Nachrichten und die InfluxDB hat der Edge-KI Rechner Zugriff auf die Sensorwerte und Zustandsgrößen des technischen Prozesses und kann damit den Zuckergehalt über eine künstliche Intelligenz (KI) bestimmen. Auf oberster Ebene wird den Mitarbeitern über ein System zur Betriebsdatenerfassung (BDE) eine grafische Oberfläche (GUI) angeboten, über die eine Steuerung des Prozesses und eine Interaktion mit den übrigen Komponenten möglich wird.

### Detaillierte Beschreibung des Gesamtsystems

#### Virtuelle Sensoren

Virtuelle Sensoren (VISO) bestehen aus unterschiedlichen physikalischen Sensoren und berechnen oftmals eine ansonsten schwer zu messende Größe. Dabei ist das Messen der notwendigen Sensordaten und die Wahl der passenden Messverfahren eine der wichtigsten und aufwendigsten Aufgaben. Im Abschnitt Motivation wurde bereits ein Beispiel erwähnt, in dem Ultraschall eingesetzt wird, um die Dichte einer Flüssigkeit zu messen um dann erst auf den Zuckergehalt zu schließen. Somit handelt es sich auch hier bereits um einen VISO, da die Zielgröße nicht direkt gemessen, sondern über verschiedene Sensoren und Algorithmen indirekt ermittelt wird. Zur Berechnung wird entweder ein physikalisches Modell,

ein analytisches Modell oder eine Kombination aus beiden eingesetzt. Diese Modelle sollen die physikalischen Zusammenhänge abbilden und anhand der Messwerte zuverlässige Ergebnisse liefern.

Physikalische Modelle werden hauptsächlich eingesetzt, wenn die Zusammenhänge der verschiedenen Zustandsgrößen bekannt sind und alle notwendigen Größen einfach gemessen werden können.

Analytische Modelle kommen oft zum Einsatz, wenn der genaue physikalische Zusammenhang noch nicht komplett bekannt ist oder bestimmte Größen nicht mit gewöhnlichen Sensoren gemessen werden können. In diesen Fällen kann dann eine KI eingesetzt werden, um aus den vorhandenen Daten eine Beziehung zur Zielgröße zu bestimmen und diese Vorhersage durch neue Daten stetig zu verbessern.

In der Regelungstechnik wird für diesen Zweck ein sogenannter Beobachter (als Beispiel für ein physikalisches Modell) verwendet [1]. Für eine genaue und passende Regelung werden verschiedene Zustandsgrößen benötigt, die aber bei vielen Systemen nur unter großem Aufwand messbar sind. Der Beobachter schafft es hier mit Hilfe eines mathematischen Modells (Differenzialgleichungen) die Zustandsgrößen rechnerisch zu bestimmen. Damit das Modell eine ausreichende Genauigkeit erreicht, werden Sensorwerte aus dem realen Prozess eingespielt. David Luenberger entwickelte dafür den Luenberger-Beobachter [2]. Ähnlich zum Luenberger-Beobachter beschreibt Gazaille in [3] einen Soft-Sensor um Größen aus einem Bioprozess zu ermitteln. Dieser besteht sowohl aus einem realen Sensor, als auch aus einem Schätzer, um die gewünschte Größe zu ermitteln.

Wolfgang Samenring [4] patentierte ein Verfahren, um die Qualität, insbesondere die Viskosität, von Öl bei Verbrennungsmotoren zu bestimmen. Er zeigte über Schaubilder einen Zusammenhang zwischen der Viskosität und der Verunreinigung. Dabei werden bei dem vorgeschlagenen Verfahren lediglich die Temperatur und der Öldruck gemessen. Beides sind Größen, die sich mit herkömmlichen und einfachen Sensoren bestimmen lassen. So lässt sich über eine mathematische Funktion mit Temperatur und Öldruck als Parameter, die Viskosität bestimmen. Ein weiterer interessanter Aspekt ist der Einsatz von Kontextdaten, bspw. die Zeitpunkte der Serviceintervalle und des Ölwechsels. Diese werden berücksichtigt und beeinflussen das Datenmodell. Dieser Ansatz fließt mit in das Projekt ein und wird weiterverfolgt, um bspw. Unterschiede in den verwendeten Rohstoffen festzustellen und diese Erkenntnisse an das Modell weiterzugeben.

#### *Rolle des Edge-KI Rechners für den virtuellen Sensor*

Viele physikalische Größen können mit herkömmlichen und einfachen Sensoren nur bestimmt werden, wenn die physikalischen Zusammenhänge des technischen Prozesses bekannt sind. Ein bereits erwähntes Beispiel ist die Bestimmung des Zuckergehalts in Bierwürze. Ein einfaches Messprinzip dabei ist die Bestimmung der Dichte mit zwei Drucksensoren, die mit einem bekannten Höhenunterschied verbaut sind (siehe Abbildung 2). Drucksensoren sind auf dem Markt gut erhältlich und können

ohne allzu großen Aufwand in den Braubehälter installiert werden. Die Dichte kann jedoch nicht direkt über den hydrostatischen Druck bestimmt werden, da weitere Parameter in die Dichteberechnung miteinfließen. Ein Beispiel ist die geometrische Anordnung der Drucksensoren, insbesondere die Entfernung zur Heizspirale im Bierbraukessel. Auch kann die Temperaturverteilung innerhalb der Flüssigkeit variieren, wodurch sich deutliche Schwankungen in der Dichtemessung ergeben können. Dies wurde in den ersten Vorversuchen des Projektes als Problem identifiziert und durch eine geschickte Anordnung der Sensoren weitestgehend gelöst.



*Abbildung 2: Zwei Ein-Zoll-Einschraubstutzen für Drucksensoren*

Trotzdem ist es wichtig nicht nur die Sensorwerte (Druck, Temperatur), sondern auch weitere Zustandsgrößen (Heizspirale an/aus) zu erfassen. Dadurch wird die Modellbildung mit einem physikalischen Modell jedoch schwierig, da die Gesamtheit aller Zustandsgrößen das physikalische Modell sehr komplex gestalten würde. Aus diesem Grund wird vorgeschlagen, das physikalische Modell durch eine KI zu ersetzen.

Oftmals sind herkömmliche Rechner für diese Aufgaben nur bedingt geeignet, da das Trainieren der KI sehr rechenintensiv sein kann. Dabei sind vor allem eine leistungsfähige CPU oder eine für KI geeignete GPU sehr wichtig. Das Trainieren der KI wird daher auf einen Rechner, im Weiteren Edge-KI genannt, mit dedizierter GPU ausgelagert.

#### *Sensorwerte aus einem Wassertest-Experiment zur Validierung der Dichtemessung*

Um das Messverfahren zu validieren, wurde ein Versuch ohne Malz durchgeführt. Abbildung 3 zeigt den Temperaturverlauf des Wassers im Braukessel. Zunächst steigt die Temperatur auf 63°C und wird dort für ca. 70 Minuten gehalten. Diese Phase nennt man in der Brautechnik die erste Rast (rot). Das Halten der Temperatur wird über einen Zweipunktregler der SPS gesteuert. Danach erfolgt

ein weiterer Temperaturanstieg auf 73°C und das Halten der Temperatur für ca. 5 Minuten. Dieser Abschnitt wird als zweite Rast (blau) bezeichnet. Danach erfolgt mit dem Abmaischen (grün), im Diagramm und im Weiteren als Rast 3 bezeichnet, der letzte Temperaturanstieg auf 78°C mit einer Haltedauer von 5 Minuten.



Abbildung 3: Temperaturverlauf von Wasser

Abbildung 4 veranschaulicht den Verlauf der Dichte, welche durch die zwei Drucksensoren berechnet wurde. Aufgrund des Temperaturanstiegs, sinkt auch die Dichte des Wassers (orange). Sobald die Temperatur von der SPS gehalten werden wird, bleibt die Dichte des Wassers im Mittel konstant (rot). Beim Aufheizen zur zweiten Rast nimmt die Dichte des Wassers weiter ab (gelb), bis die zweite Rast erreicht wird (blau). Beim Aufheizen zur dritten Rast nimmt die Dichte des Wassers nochmals ab (lila) und wird wieder konstant, solange die Rast anhält (grün). Der dazugehörige Temperaturverlauf ist in Abbildung 3 abgebildet.



Abbildung 4: Verlauf der Dichte

### Sensorwerte aus einem Maische-Prozess zur Validierung der Dichtemessung

Im Folgenden wird auf die Sensorwerte während des Brauprozesses eingegangen, insbesondere auf die Werte der zwei Drucksensoren. Hierbei wird der Bierbraukessel mit Wasser gefüllt und auf 40°C aufgeheizt. Dann wird Malz zugegeben und der Inhalt auf 63°C erhitzt und diese Temperatur für ca. 70 Minuten gehalten. Abbildung 5 zeigt diesen Temperaturverlauf. Ziel dieser ersten Rast ist es, einen biochemischen Prozess in Gang zu setzen. Dabei wandeln im Malz enthaltene Enzyme ( $\beta$ -Amylase) die Stärkemoleküle des Malzes hauptsächlich in kurzketige und vergärbare Zucker (Maltose) um.



Abbildung 5: Temperaturverlauf von Maische

Abbildung 6 zeigt den geglätteten Dichteverlauf (grün), berechnet aus der Differenz der beiden Drucksensoren. Die Druckdifferenz nimmt während der ersten Rast stets zu. Weitere Verläufe in Abbildung 6 sind die manuellen Zuckermessungen mit der Bierspindel. Hier werden Proben aus dem Bierkessel entnommen und die Bierspindelwerte direkt nach Entnahme (unkalibriert, hellblau) und nach dem Aufheizen (kalibriert, orange) aufgezeichnet. Hier ist zu sehen, dass die Werte der Dichte und der manuellen Messungen ähnliche Verläufe haben. Nach 70 Minuten geht der Prozess in seine nächste Phase, die zweite Rast. Es gibt hier ein Aufheizen auf 73°C und ein Halten für 5 Minuten. Dabei werden verbleibende Stärkemoleküle durch die Enzyme ( $\alpha$ -Amylase) in größtenteils langkettige, nicht vergärbare Zucker (Dextrin) umgewandelt. Durch das Aufheizen der Flüssigkeit ist in Abbildung 6 auch eine Abnahme der Dichte ersichtlich. Nach Ende der zweiten Rast gibt es ein letztes Aufheizen auf 78°C. Bei dieser Rast werden die Enzyme und Proteine durch die Hitze zerstört und somit der biochemische Prozess zunächst beendet. Die letzte Phase des Prozesses wird Läutern genannt. Hier wird das Malz aus der Bierwürze gehoben und anschließend Wasser über den Trester (Überbleibsel vom Malz) gegossen. Das Wasser löst den noch im Trester enthaltenen Zucker. Bei den manuellen Messungen ist hierbei eine Abnahme der Bierspindelwerte zu beobachten.



Abbildung 6: Verlauf der Dichte und Zuckergehalts

### Datenmanagement

Für das Trainieren des virtuellen Sensors wird eine große Datenmenge benötigt. Dabei ist zunächst zu bestimmen, welche Daten im Prozess eine besondere Bedeutung haben. Dies erfolgt i.d.R. durch eine Kombination aus Expertenwissen und experimentellen Analysen. Letzteres kann zum einen durch das automatisierte Aufzeichnen von Daten bei der Prozessausführung geschehen. Zum anderen können aktiv Experimente für spezifische Situationen geplant werden. Der zweite Ansatz ist insbesondere dann sinnvoll, wenn lediglich Teile des Gesamtprozesses für das Training der Modelle relevant sind. Der Gesamtprozess zur Herstellung von Bier beinhaltet nach der Herstellung der Bierwürze noch weitere Schritte, z.B. das Hopfenkochen oder Reifen. Diese Schritte sind aber für die Datengewinnung völlig unerheblich, weshalb auf eine Weiterverarbeitung der Würze verzichtet wurde. Neben der erwarteten Zeitersparnis war bei der Entscheidung über die Datengewinnung die größere Flexibilität bei der Variation der Trainingsbedingungen zentral. Für die Generierung valider Trainingsdaten wurden drei wesentliche Anforderungen definiert:

1. Größtmögliche Transparenz über den Prozess und das Zustandekommen der Daten
2. Vermeidung von Fehlern bei der Versuchsdurchführung
3. Berücksichtigung von Kontextdaten, bspw. Prozessinformationen, Beschaffenheit der Rohstoffe etc.

Aus den Überlegungen zur letzteren Anforderung entstand die Annahme, dass Umweltbedingungen (bspw. Umgebungstemperatur und Luftfeuchtigkeit) oder die Beschaffenheit der Ausgangsstoffe Einfluss auf die Versuche haben können. Dies ist auch unmittelbar einleuchtend,

| Vorgangsnummer | Arbeitsplatz | Werk | Steuerschlüssel | Beschreibung                           | Rüstzeit | Maschinenzeit | Personalzeit | Rüsten variable |
|----------------|--------------|------|-----------------|--|----------|---------------|--------------|-----------------|
| 0005           | P01-001      | DF01 | Z001            | PH-Messer kalibrieren (manuell)        | 5        | 0             | 0            | 0               |
| 0010           | P01-001      | DF01 | Z002            | Wasser einfüllen (manuell)             | 5        | 0             | 0            | 0               |
| 0015           | P01-001      | DF01 | Z004            | Temperieren                            | 0        | 0             | 0            | 5               |
| 0020           | P01-001      | DF01 | Z002            | PH-Wert messen und anpassen (manuell)  | 3        | 0             | 0            | 0               |
| 0025           | P01-001      | DF01 | Z002            | Malz zugeben (manuell)                 | 2        | 0             | 0            | 0               |
| 0030           | P01-001      | DF01 | Z004            | Heizen auf Zieltemperatur              | 0        | 0             | 0            | 5               |
| 0035           | P01-001      | DF01 | Z004            | Rast 1                                 | 0        | 70            | 0            | 0               |
| 0045           | P01-001      | DF01 | Z004            | Heizen auf Zieltemperatur              | 0        | 0             | 0            | 5               |
| 0050           | P01-001      | DF01 | Z004            | Rast 2                                 | 0        | 5             | 0            | 0               |
| 0065           | P01-001      | DF01 | Z004            | Heizen auf Zieltemperatur              | 0        | 0             | 0            | 5               |
| 0075           | P01-001      | DF01 | Z004            | Rast 3                                 | 0        | 5             | 0            | 0               |
| 0085           | P01-001      | DF01 | Z001            | Treber heben (manuell)                 | 5        | 0             | 0            | 0               |
| 0090           | P01-001      | DF01 | Z002            | Zuckergehalt messen (manuell)          | 3        | 0             | 0            | 0               |
| 0095           | P01-001      | DF01 | Z002            | Läutern (manuell)                      | 10       | 0             | 0            | 0               |
| 0100           | P01-001      | DF01 | Z005            | Mengenfeststellung / Auftragsabschluss | 1        | 0             | 0            | 0               |

Tabelle 1: Arbeitsplan für einen Brauversuch

tend, da erstens bei den Rohstoffen unterschiedliche Sorten verwendet werden und zweitens selbst bei der gleichen Sorte nur bedingt kontrollierbare Unterschiede in deren Beschaffenheit wie z.B. Mehlgelbheit, Wassergehalt etc. zwischen zwei Losen vorkommen können. Die chemisch-physikalischen Eigenschaften des Malzes liegen der Lieferung jeweils in Form eines Laborberichts bei und können so für das Training der KI verwendet werden. Hierfür sind eine entsprechende Erfassung, Verwaltung und Bereitstellung der Daten notwendig.

Beim Sammeln von Trainingsdaten gibt es verschiedene Fehlerquellen. Eine Quelle von Fehlern besteht in der nicht sachgemäßen Ausführung der Versuche, also z.B. dem Überspringen oder Auslassen von Prozessschritten, der Missachtung von Prozessvorgaben etc. Somit muss der Versuchsaufbau diese Einflüsse so gut es geht auf ein Mindestmaß reduzieren und gegebenenfalls bei nicht konformer Prozessausführung dies entsprechend vermerken. Eine weitere Fehlerquelle ist die manuelle Erfassung von Messergebnissen. Diese falschen Messergebnisse können zum einen durch falsches Ablesen der Messinstrumente oder zum anderen durch Eingabefehler bei der Messwertaufzeichnung entstehen. Auch ein einfaches Auslassen von einzelnen Messintervallen ist möglich, wenn der Labormitarbeiter die Durchführung der Messung vergisst.

Abschließend ist die Nachvollziehbarkeit der erfassten Daten und die Möglichkeit der Prozessüberwachung gerade bei unterschiedlichen Rezepten ohne eine systematische Verwaltung dieser Daten über ein betriebliches Informationssystem (bIS) nur schwer umsetzbar.

Im Bereich bIS besteht bei den Projektteilnehmern eine langjährige Erfahrung mit ERP-Systemen, speziell mit

dem Produkt S/4HANA von SAP. Dieses steht der Hochschule Albstadt-Sigmaringen als eigenes System zu Ausbildungs- und Forschungszwecken zur Verfügung. Teile der Anforderungen können bereits mit der Standardauslieferung der Systems abgedeckt werden, weitere Funktionen können aufgrund der Flexibilität des Systems mit vertretbarem Aufwand durch Systemkonfigurationen und Code Anpassungen hinzugefügt werden.

Durch das große Ökosystem an verfügbaren Erweiterungen und Add-Ons, rund um S/4HANA, kann die Standardsoftware an spezifische Rahmenbedingungen angepasst werden.

Die Führung der Labormitarbeiter durch den Versuch ist ein zentraler Baustein des bIS. Zunächst müssen hierfür

sämtliche relevanten Daten identifiziert und eine passende SAP Darstellung gewählt werden. Im Wesentlichen werden die Produktstruktur, also die Ausgangsstoffe mit entsprechenden Mengenangaben, sowie der konkrete Ablauf des Versuchs mit diesen Daten beschrieben. Jedem Vorgang können Attribute beziehungsweise Merkmale zugeordnet werden, die entweder Vorgaben für die Prozessautomatisierung darstellen oder für die Messdatenerfassung verwendet werden.

Die Realisierung in SAP erfolgt mittels Arbeitsplänen, siehe Tabelle 1, und den einzelnen Vorgängen zugeordneten Prüfmerkmalen. Die Prüfmerkmale können sowohl Prozessvorgaben als auch zu erfassende Messgrößen darstellen. Die Unterscheidung erfolgt über spezifische Parameter in den jeweiligen Stammprüfmerkmalen. Die Erfassung der Prüfergebnisse wird in SAP mittels Prüfpunkten umgesetzt, so dass seine den Versuch begleitende Prüfung ermöglicht wird.

Wie bereits erwähnt, gibt es natürliche Schwankungen in der Zusammensetzung des Hauptrohstoffs Malz, so dass diese Informationen zu jedem Sack des Malzes gespeichert werden müssen. Dies erfordert eine konsequente Identifikation der Malzsäcke, die als Charge im bIS verwaltet werden, sowie eine konsistente Erfassung der in den Versuchen eingesetzten Chargen. SAP bietet für dieses Szenario mit der Chargenklassifikation ein sehr einfach zu verwendendes aber auch mächtiges Instrument zur Beschreibung von Chargeneigenschaften. Relevante Daten, wie Wasser-, Mehlgelbheit etc. werden aus dem Begleitpapier (Laborbericht) als Chargenmerkmal gespeichert, siehe Tabelle 2. Die manuelle Eingabe der Daten

kann durch den Einsatz einer SAP-Erweiterung, die in einem parallel durchgeführten Studierendenprojekt entstand, und auf Optical Character Recognition (OCR) aufbaut, einfach automatisiert werden. Hierdurch kann eine hohe Datenqualität erreicht werden und jede Anlieferung führt direkt zu einem Wareneingang mit einer neuen, automatisch klassifizierten Charge.

| Parameter                | Einheit | Messwert |
|--------------------------|---------|----------|
| Wassergehalt             | %       | 3,5      |
| Extrakt Mehl lufttrocken | %       | 79,4     |
| Extrakt Mehl wasserfrei  | %       | 82,2     |
| Sortierung 2,8mm         | %       | 81,9     |
| Etc.                     | Etc.    | Etc.     |

Tabelle 2: Ausschnitt aus Begleitpapier (Laborbericht)

Jeder Versuch wird über einen eigenen Auftrag im bis abgebildet. Zum Auftrag werden zunächst die eingesetzten Rohstoffe, deren Menge und die ausgewählte Charge

| Sensordaten von SPS empfangen, alle 10 Sekunden |     |  | Sensorname         | Datentyp        | Einheit | Anmerkung   |
|---|-----|--|--------------------|-----------------|---------|---|
| Topic   | QoS | Mqtt Payload   |                    |                 |         |   |
| sps/edge/values                                 | 2   | {"timestamp":<Long>,"operation":<String>,"batchid":<String>,"emergencyShutdown":<bool>,"hatch":<bool>,"sensorValues":{"pressure_high":20.7,"pressure_low":31.65,"temperature":68.3,"weight":34.8,"<String>":<Float>...}} | temperature        | Float           | °C      |   |
|   |     |  | weight             | Float           | kg      | -100 wenn Waage aus ist                                 |
|   |     |  | pressure_high      | Float           | mbar    |   |
|   |     |  | pressure_low       | Float           | mbar    |   |
|   |     |  | power_stir         | Float           | mA      |   |
|   |     |  | voltage_stir       | Float           | V       |   |
|   |     |  | power_heat         | Float           | mA      |   |
|   |     |  | voltage_heat       | Float           | V       |   |
|   |     |  | emergency_shutdown | Boolean         |         | True ist der kritische Fall, False ist der normale Fall |
|   |     |  | hatch              | Boolean         |         | False nicht drauf und True drauf                        |
|   |     |  | timestamp          | EpochTime(Long) |         |   |

Tabelle 3: Beispiel für eine MQTT-Nachricht (Toggle-Nachricht)

erfasst und über eine Bestandsbuchung, also einer Warenentnahme, dem Auftrag zugeordnet.

Der Auftrag im bis stellt somit das zentrale Element bei der Bereitstellung der Prozessvorgaben, Kontextdaten sowie für die Durchführung der Messungen dar. Diese Daten müssen auch für die anderen Komponenten des Gesamtsystems verfügbar sein. Grundsätzlich stehen für den systemübergreifenden Austausch von Daten ins S/HANA verschiedene Ansätze zur Wahl, bspw. IDoc, RFC, WebServices, OData-Service etc.

Bereits in zahlreichen Projekten hat sich die Kommunikation über das im Internet of Things (IoT) Bereich weitverbreitete MQTT-Protokoll bewährt.<sup>1</sup>

Waren in der Vergangenheit für diese Kommunikation noch separate Komponenten wie z.B. SAP JCo, SAP PCo oder andere Middleware, zwischen SAP und der übrigen Infrastruktur notwendig, so kann diese nun direkt integriert über eine SAP Klasse erfolgen. Ein permanentes Überwachen der abonnierten Topics wird mit einem SAP-Daemon realisiert.

Gemeinsam mit dem Projektpartner top flow wurde eine in das ERP-System vollständig integrierte MES-Lösung implementiert. In großen Teilen konnte auf Standardkomponenten der Firma top flow zurückgegriffen werden und durch das flexible Framework auf die Projektspezifika angepasst werden. Das Kernstück der Erweiterung besteht aus einer BDE-Anwendung, über das Labormit-

arbeiter alle relevanten Daten angezeigt und manuelle Messdaten eingetragen werden können.

## Systemarchitektur und -betrieb

Der schematische Aufbau des Gesamtsystems wurde bereits vorgestellt. Im Folgenden soll nun näher auf die Orchestrierung der einzelnen Komponenten eingegangen werden.

Grundlage für die Kommunikation sämtlicher Teile des Gesamtsystems ist das MQTT-Protokoll. Über einen Broker können die Teilsysteme auf verschiedenen Topics miteinander kommunizieren. Ein Beispiel für eine der Nachrichten ist in Tabelle 3 zu sehen. Zur Benennung der Topics wurde die Konvention "Sender/Empfänger/ Bedeutung" verwendet. An dem Topic edge/sps/toggle erkennt man bspw., dass es sich um eine Toggle-Nachricht

handelt, die vom Edge-Server verschickt und von der SPS, bzw. vom Gateway, verarbeitet wird.

Zur Beschreibung der Datenflüsse im Rahmen der Prozessausführung wurde ein vereinfachtes Sequenzdiagramm verwendet, das in Abbildung 7 dargestellt ist.

Aus der Darstellung erkennt man, dass die Informationen über den Edge-Server in zwei Richtungen weitergeleitet werden. Zum einen werden die Daten in einer InfluxDB-Datenbank gespeichert und dort für das Training der neuronalen Netze bereitgestellt. Zum anderen werden bestimmte Daten an die SPS durchgeschleift. Zusätzlich wird auf dem Edge-Server der aktuelle Prozesszustand gespeichert und verwaltet.

Während der Design- und Testphase erwiesen sich im Hinblick auf die Prozessausführung immer wieder Ausfälle einzelner Komponenten im Gesamtsystem als problematisch. So war bspw. zunächst beim Ausfall des MQTT-Brokers nicht gewährleistet, dass keine Daten verloren gehen oder bei einem Ausfall des Edge-Servers die erfassten Daten dennoch in die Datenbank geschrieben werden. Eine Risikoanalyse ergab, dass insbesondere die folgenden Fehlerfälle ein hohes Risiko für Daten- bzw. Konsistenzverlust innerhalb der Prozessausführung darstellen und aus diesen Gründen eine Risikovermeidung notwendig ist. Eine Mindmap über verschiedene Risikoszenarien ist in Abbildung 8 dargestellt.

<sup>1</sup> Als Middleware fungiert eine JAVA-Application. Diese subscribt sich als MQTT-Client auf die zur Anlagensteuerung relevanten Topics, verarbeitet und interpretiert die Daten und löst schließlich Aktionen im

ERP-System aus. Konkret erfolgt unter Verwendung des Java-Connectors der Aufruf bestimmter RFC-fähiger Bausteine im S/4HANA System.

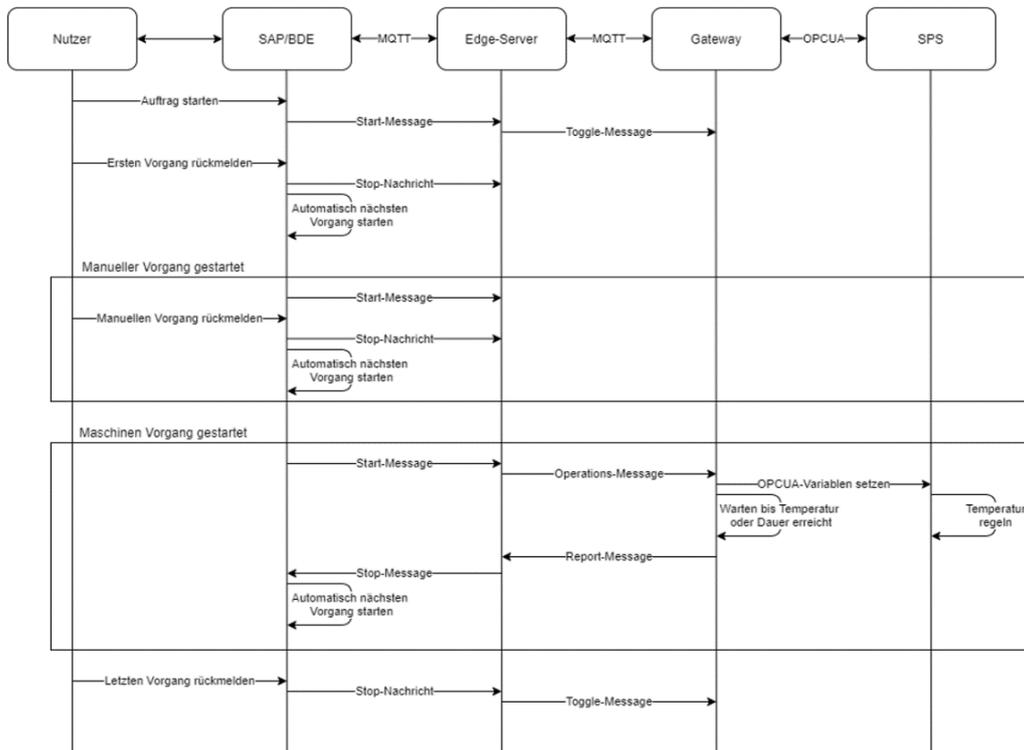


Abbildung 7: Sequenzdiagramm

Folgende Hauptfehlerquellen wurden identifiziert:

- Absturz Gateway
- Absturz State-Machine
- Absturz MQTT-Broker
- Absturz SAP MQTT-Daemon

Der erneute Start nach dem Absturz einer der obigen Komponenten erfordert komponentenspezifische Lösungen, welche im Folgenden näher erläutert werden.

Um den grundsätzlichen Ausfall von Komponenten feststellen zu können, verschicken SAP und Edge-Server in regelmäßigen Abständen eine Alive-Nachricht. Für das Gateway übernimmt die Nachricht zur Übermittlung der Sensorwerte diese Funktion, da diese bereits in regelmäßigen Abständen versendet wird.

Beim Gateway werden einige Zustandsinformationen (Zeitstempel, Dauer, Temperatur und Status) beim Empfangen auf die SPS geschrieben und dort gespeichert. Beim Neustart werden diese Informationen eingelesen und von der SPS beim Restart verwendet.

Nach dem Absturz wird der Edge-Server automatisch neu gestartet. Der aktuelle Zustand kann anhand von SAP Alive-Nachrichten ermittelt und die State-Machine wieder in den passenden Zustand gebracht werden.

Damit bei einem Ausfall des Internets oder des zentralen MQTT-Brokers trotzdem Daten von den Sensoren nicht verloren gehen, wurde ein lokaler MQTT-Broker installiert, der die Nachrichten im Fehlerfall zwischenspeichert. Sobald die Verbindung zum zentralen MQTT-Broker wieder hergestellt ist, werden die Daten an diesen weiter-

geleitet und in die zentrale Datenbank geschrieben. Zusätzlich werden die Daten lokal auf dem Gateway in eine separate InfluxDB geschrieben, wodurch selbst bei einem Internetausfall die Daten weiterhin vorhanden sind oder bei Übertragungs- oder Nutzerfehlern ein Backup vorhanden ist. In der Versuchsdurchführung hat sich dieses Vorgehen als sinnvoll erwiesen, da sowohl ein Absturz der lokalen Datenbank als auch ein Internetausfall auftraten, dabei aber keine Prozessdaten verloren gingen. Auf Seiten von SAP wurde ein Service eingerichtet, der den MQTT-Daemon überwacht und diesen neu startet, falls die Verbindung zum Broker verloren geht.

### Modellbildung und Training

Sämtliche Sensordaten aus dem Bierbrauprozess werden in einer InfluxDB-Datenbank gespeichert. InfluxDB ist eine Zeitreihen-basierte Datenbank, deren Fokus auf der Verarbeitung von Daten mit Zeitinformationen liegt. Aus diesem Grund ist sie sehr gut geeignet, um bspw. Sensordaten zu speichern und diese aufzubereiten. Es werden jeweils von sieben Sensoren Werte gesammelt: Temperatursensor, Drucksensor-oben, Drucksensor-unten, Deckel-Sensor, Leistung-Heizung, Leistung-Rührer und Gewicht. Aus diesen Sensorwerten können weitere Daten abgeleitet werden, bspw. enthält die Leistung der Heizung Informationen über deren aktuellen Zustand (Heizung an oder aus). Das Gateway sendet etwa alle zehn Sekunden eine Nachricht mit den aktuellen Sensordaten. Neben diesen beobachtenden Daten gibt es auch manu-

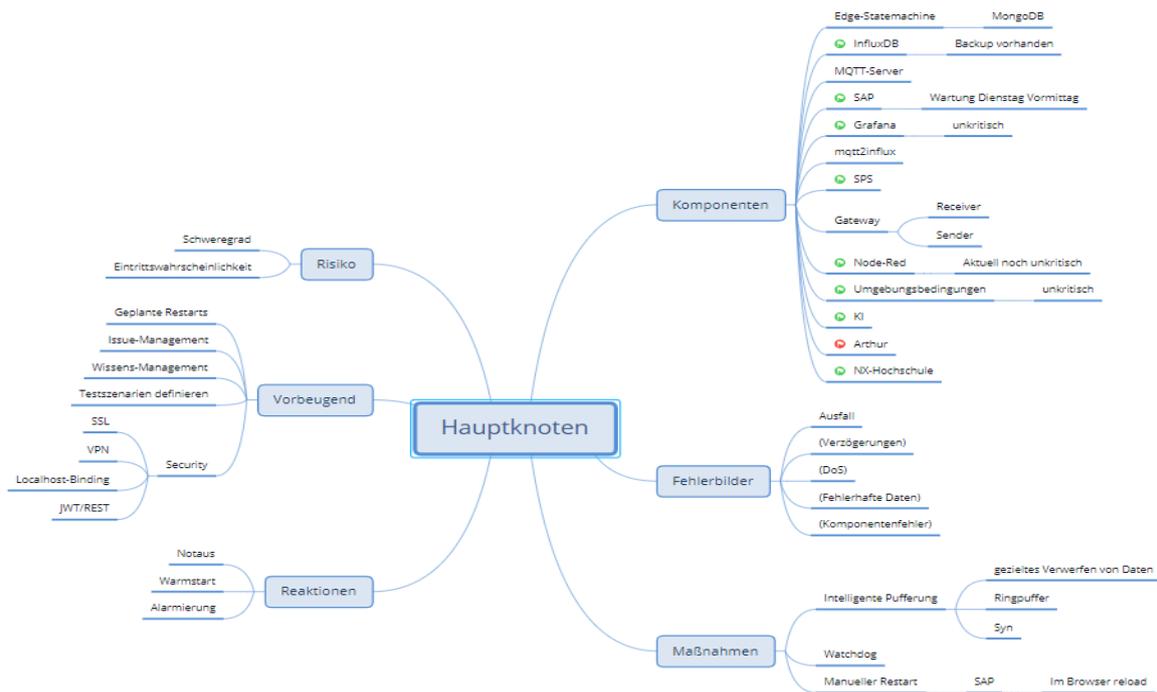


Abbildung 8: Mindmap zur Risikobewertung

elle Messungen. Dazu zählt unter anderem das Ziel-Feature (Target-Feature), der Bierspindelwert. Weitere manuelle Messungen sind bspw. der pH-Wert und die Messung des Bierspindelwerts bei einer auf 70°C aufgeheizten Probe. Diese Messungen benötigen einen manuellen Eingriff in den Prozess und werden etwa alle zehn Minuten getätigt, über das bIS eingetragen und in die Datenbank übernommen. Aufgrund der deutlich geringeren Frequenz der manuellen Messungen müssen diese Datensätze vor dem Training noch aufbereitet werden. Hierfür werden die Daten unter anderem interpoliert, um jedem Zeitpunkt einen Zielwert zuweisen zu können. Die Ergebnisse der Aufbereitung werden in eine CSV-Datei gespeichert, um die weitere Verarbeitung zu erleichtern und Zeit für zusätzliche InfluxDB-Queries zu sparen.

Während des Trainierens des Netzes hat es sich als vorteilhaft erwiesen, bestimmte Messwerte nicht als direktes Eingangs-Feature zu verwenden. Beispielsweise haben Stufenfunktionen einen sehr starken Einfluss auf die Vorhersagen des Netzes, da bei der zeitlichen Ableitung durch die Sprünge sehr große Werte entstehen, die dann einen ebenso großen Einfluss auf die Optimierung der Gewichte haben. Diskrete Zustandsgrößen werden jedoch als Hilfsgrößen verwendet, um bspw. das Öffnen und Schließen des Bierbraudeckels zu dokumentieren und Messungen, die zu diesen Zeitpunkten erfolgten herauszufiltern. Aus den gesammelten Sensorwerten werden im aktuellen Modell nur noch drei verwendet, die Temperatur, der obere Druck sowie die Druckdifferenz. Als Zielgröße (Target-Feature) dient der unkalibrierte Zuckermesswert.

Für zeitbasierte Daten werden oft LSTM-Modelle (Long-Short-Term-Memory) verwendet, um den zeitlichen Zusammenhang der Daten zu berücksichtigen. Diese zeichnen sich durch das Speichern bestimmter Zustandsgrößen (Hidden und Cell State) aus, wodurch zeitliche und dynamische Aspekte durch die Daten einen Einfluss auf die aktuelle Berechnung haben. Während des Trainings werden die Gewichte des Modells so angepasst, dass eine Fehlerfunktion minimiert und damit die Vorhersagen sich immer mehr den gemessenen Werten annähern. Das verwendete Modell besteht aus drei Schichten. Die erste Schicht besteht aus 32 LSTM Einheiten, darauf folgt eine Dense-Schicht mit 64 Einheiten und eine letzte Dense-Schicht mit nur einem Neuron, siehe Abbildung 9.

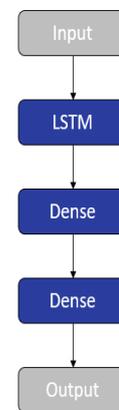


Abbildung 9: verwendetes LSTM

Die Verbindungen zwischen den einzelnen Schichten bzw. Einheiten werden Gewichte genannt und werden im Trainingsverlauf optimiert. Das gewählte Netz ist relativ

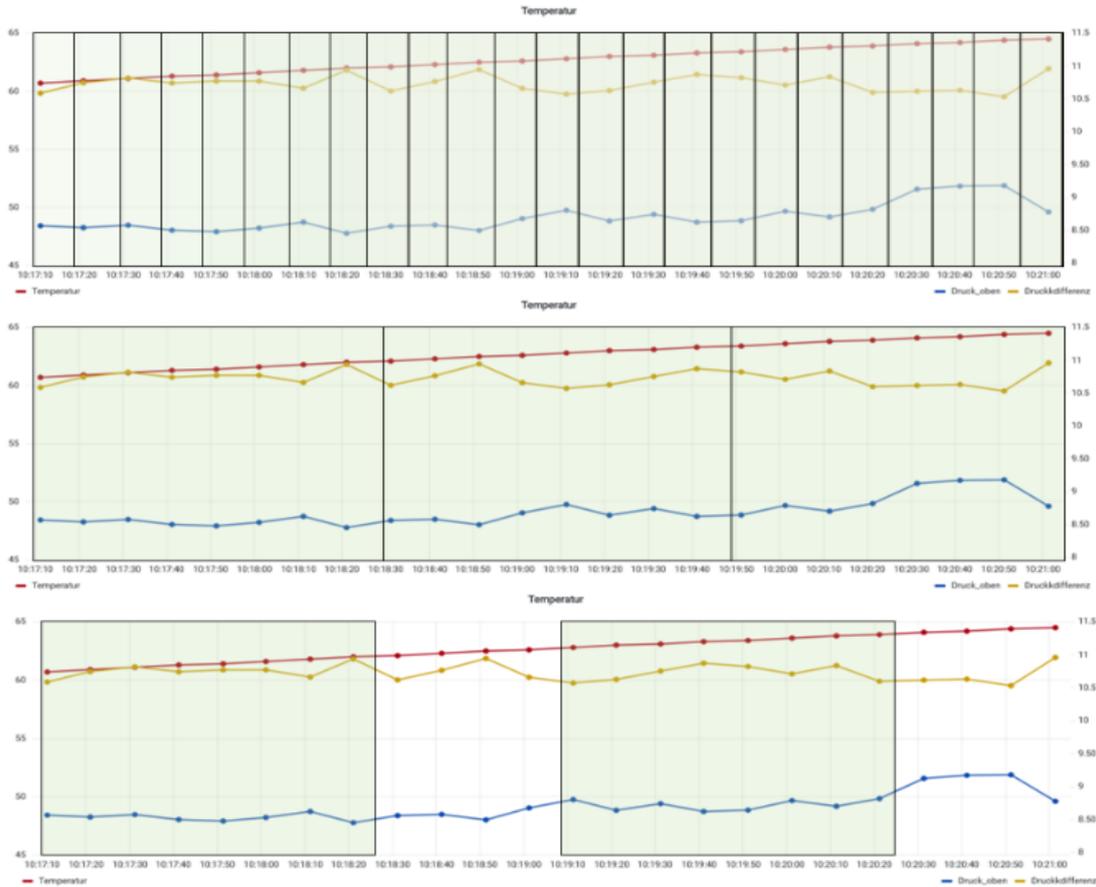


Abbildung 10: Optionen für die Festlegung von Zeitfenstern

klein, wobei beim Training auch größere Netze getestet wurden, durch die allerdings keine Verbesserung bzgl. der Vorhersagegüte erzielt werden konnte. Durch das kleine Netz können sehr flexibel und schnell verschiedene Trainingseinstellungen (Hyperparameter) verändert und getestet werden.

Bevor das Netz die Daten erhält, werden die Daten normalisiert, also in einen bestimmten Wertebereich transformiert, damit große und kleine Datenwerte den gleichen Einfluss haben. Zum Normalisieren wurde der Z-Score [5] verwendet.

Während des Trainings liest der Trainingsalgorithmus acht zeitlich aufeinanderfolgende Zeitpunkte mit den jeweiligen Attributen aus und fügt sie zu einer  $8 \times 3$  Matrix zusammen. Die entstandene Matrix kann man sich als Zeitfenster im Diagramm vorstellen. Nun hat man mehrere Optionen, dieses Zeitfenster im nächsten Schritt zu verschieben (siehe Abbildung 10). Option eins (obere Graphik) ist das Zeitfenster um eine Zeiteinheit nach vorne zu schieben, sodass sich das neue und das alte Zeitfenster überschneiden. Eine zweite Option (mittlere Darstellung) ist, genau die Länge des Zeitfensters nach vorne zu gehen, so dass die Zeitfenster aneinander angrenzen, sich aber nicht überschneiden. Eine dritte Option (untere Graphik) ist es, einen größeren Schritt zu machen, so dass eine Lücke zwischen zwei Zeitfenstern entsteht, hier würden allerdings weniger Trainingsdaten entstehen.

Für das Netz wurden Optionen eins und zwei getestet. Im Training hat sich gezeigt, dass Option zwei, aneinander angrenzende, nicht überlappende Zeitfenster, bessere Ergebnisse liefern und die Modelle schneller konvergieren. Auf Option drei wurde aufgrund der sich verringernenden Datenmenge verzichtet.

Abbildungen 11 und 12 zeigen Versuche mit Testdaten an denen die prädiktive Validität des Modells überprüft

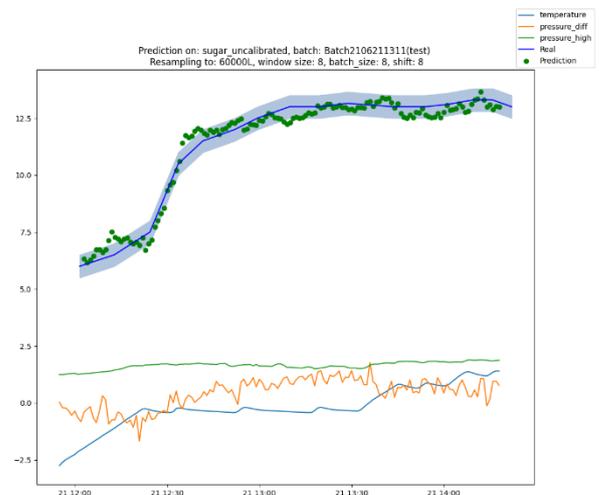


Abbildung 11: Testdaten, kleiner Fehler

wurde und bei denen das Netz sehr erfolgreich den Zuckerwert vorhersagt. Testdaten sind Daten auf die das Modell nicht trainiert wurde. Es ist klar zu erkennen, dass das Netz dem realen Verlauf folgt und die Vorhersagewerte größtenteils in dem blau eingezeichneten Toleranzbereich von 0,5 liegen.

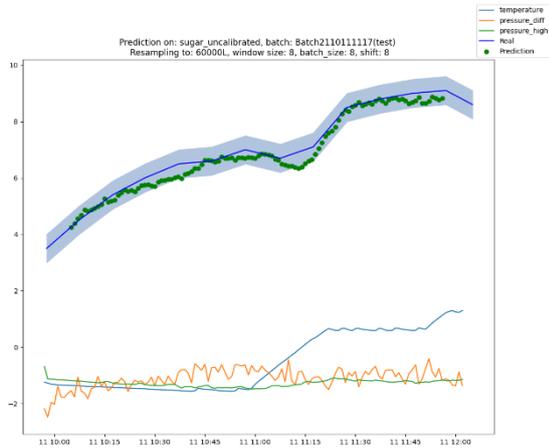


Abbildung 12: Testdaten kleiner Fehler

Bei einigen Versuchen liefert die KI jedoch auch schlechtere Ergebnisse. So ist in Abbildung 13 zu erkennen, dass die KI zunächst dem Zuckerverlauf gut folgt, dann jedoch deutlich abweicht und sich erst gegen Ende wieder den realen Werten annähert.

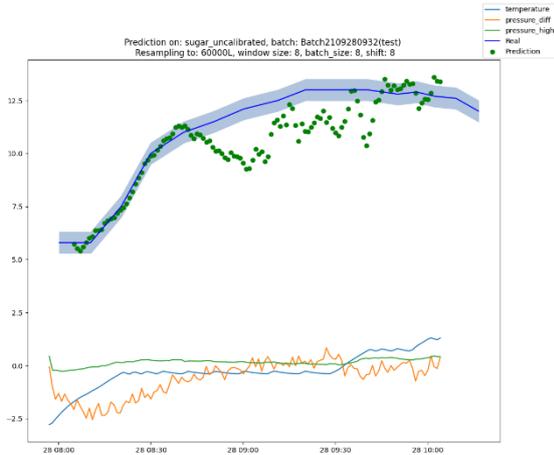


Abbildung 13: Testdaten, deutliche Abweichung

In Abbildung 14 ist ein weiterer Versuch abgebildet. Hier sind deutliche Ausreißer, die den Toleranzbereich verlassen, zu erkennen. Die KI kann grundsätzlich dem realen Verlauf trotzdem gut folgen und endet vor dem Läutern auf einem vergleichbaren Wert.

Während eines laufenden Versuchs kann die KI über Inferenz der Sensordaten live einen Vorhersagewert bilden. Dieser wird wiederum in InfluxDB geschrieben und kann anschließend live mit Hilfe von Grafana visualisiert und angezeigt werden, siehe Abbildung 15.

Das Netz braucht in diesem live Setup in fast allen Fällen weniger als eine Sekunde um das Ergebnis zu berechnen (Inferenzzeit), siehe Abbildung 16. Dadurch kann der

Vorhersagewert sehr schnell zurück in den Prozess zurückgespielt und dort weiterverarbeitet werden. Selbst die großen Ausreißer bleiben unter zwei Sekunden Inferenzzeit.

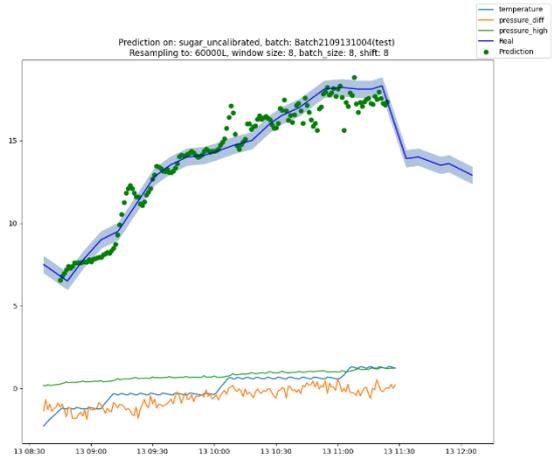


Abbildung 14: untrainierter Versuch mit Ausreißern



Abbildung 15: Vorhergesagter Zucker während eines Versuchs

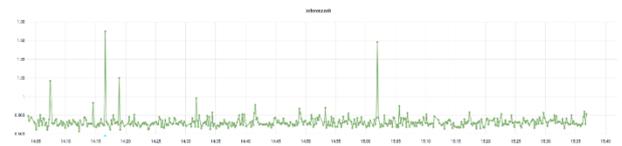


Abbildung 16: Inferenzzeit der Vorhersagewerte

## Fazit

Durch die Prozessführung durch SAP/BDE und der Darstellung aller relevanten Informationen kann der Prozess und somit die Versuche deutlich einfacher und prozesssicherer durchgeführt werden. Ein zusätzlicher Effekt ist die automatische Dokumentation von möglichen Fehlern und allgemeinen Prozess- und Kontextinformationen, wodurch eine Prozessoptimierung ermöglicht wird.

Im Projekt waren verschiedene Herausforderungen zu lösen, so hatte bspw. die anfängliche Position der Drucksensoren zu einem erheblichen Temperatur- und damit auch Dichteunterschied geführt. Eine weitere Herausforderung stellte die Resilienz des Prozesses dar. In Rasten wird die Temperatur nur für eine gewisse Dauer gehalten, ein langer Ausfall der Komponenten kann durch Überschreiten dieser Prozesszeiten zum kompletten Versuchsabbruch führen. Kürzere Ausfälle, bis etwa 10 Minuten, kann das System verkraften und dann selbstständig den Prozess fortführen.

Das Ziel, durch die KI den Zuckerwert zu bestimmen und als Prozessleitgröße zu verwenden, konnte größtenteils

umgesetzt werden. Zusätzlich können anhand der Messwerte die Vorgänge und Rezepte im bIS optimiert und angepasst werden. Beispielsweise kann im Rezept anstatt einer Dauer ein Zielwert für den Zucker angegeben und somit der Prozess flexibler gestaltet werden. Für den Bierbrauer können damit ein zu langes Heizen und damit verbundene Kosten vermieden werden.

Eine große Herausforderung in jedem KI-Projekt ist das Finden einer passenden Netzarchitektur und die passende Vorbereitung der Daten. Dazu wurden während des Projekts verschiedene Ansätze überprüft. Wobei sich grundsätzlich kleine Netze als vorteilhaft erwiesen haben, da durch das schnelle Training viele Kombinationen aus Hyperparametern und Netzarchitekturen getestet werden konnten und dies bei Ergebnissen, welche denen großer Netze ähnlich sind.

Das finale Netz liefert teilweise sehr gute Ergebnisse, teilweise aber auch noch verbesserungswürdige. Dabei ist oftmals ein Unterschied in den Eingangsdaten nur schwer erkennbar. Die Stabilität und Genauigkeit der Vorhersagen können durch das weitere Sammeln und Trainieren von Daten verbessert werden.

### Acknowledgement

Dieses Forschungsprojekt wird von verschiedenen Partnern unterstützt. Die Firma 42aaS GmbH unterstützt in Sachen KI und Infrastruktur mit ihrem Wissen und Arbeitskraft. Die Firma top flow stellt ihr Produkt zur Betriebsdatenerfassung zur Verfügung und hilft dieses an entstehende Anforderungen anzupassen. Zusätzlich hat die Firma Speidel ihr Brau-Equipment und ihre Erfahrung im Brauprozess zur Verfügung gestellt. Die Firma ifm hat verschiedene Sensoren bereitgestellt.

Dieses Projekt wurde vom Ministerium für Wirtschaft, Arbeit und Tourismus des Landes Baden-Württemberg im Rahmen ihres Programms Initiative Wirtschaft 4.0 BW gefördert und finanziert

### LITERATUR

- [1] Otto Föllinger  
Regelungstechnik, Einführung in die Methoden und ihre Anwendungen  
Erschienen: 01.06.2008  
Zugriff: 20.09.2021  
<https://www.vde-verlag.de/buecher/604201/regelungstechnik.html>
- [2] D. G. Luenberger  
Observing the state of a linear system.  
Erschienen: 01.04.1964  
Zugriff: 29.09.2021  
*IEEE Transaction on Military Electronics*. Band 8  
<https://ieeexplore.ieee.org/document/4323124>
- [3] Brian Gazaille, Benjamin Bayer  
Soft Sensors for Bioprocess Monitoring  
Erschienen: 28.05.2020  
Zugriff: 21.10.2021

<https://bioprocessintl.com/analytical/pat/soft-sensors-for-bioprocess-monitoring/>

- [4] Wolfgang Samenfink, Andreas Kufferath  
Verfahren zur Bestimmung der Ölqualität  
Erschienen: 07.12.2007  
Zugriff: 05.09.2021  
<https://patents.google.com/patent/DE102006059071A1/de>

- [5] Fred Clavel  
Basics: Standardization and the Z score  
Erschienen: 18.03.2019  
Zugriff: 18.11.2021  
<https://fredclavel.org/2019/03/18/basics-standardization-and-the-z-score/>

### KONTAKT

**Malte Hoffmann**, BS, studierte an der Hochschule Albstadt-Sigmaringen Informatik und arbeitete für ein Forschungsjahr am Projekt KI-VISOPRO.

**Derk Rembold** ist Professor an der Hochschule Albstadt-Sigmaringen und unterrichtet technische Informatik und forscht mit Schwerpunkt Cyber Physical Systems.

**Bernd Stauß** ist Professor an der Hochschule Albstadt-Sigmaringen und unterrichtet Wirtschaftsinformatik. Sein Forschungsschwerpunkt liegt im Bereich Business Analytics.

**Stefan Schwarzkopf** ist Geschäftsführer der Firma 42aaS GmbH und beschäftigt sich insbesondere mit KI Themen bei seinen Kundenprojekten.

**Marius Jacobi** studierte an der Hochschule Albstadt-Sigmaringen technische Informatik und bearbeitet seine BS Thesis bei der Firma 42aaS mit dem Thema KI.