INTERNATIONALE HOCHSCHULE **FERNSTUDIUM**

Project: Data Science Use Case (DLMDSPDSUC01)

Development and reflection phase - HVAC-Angel: Reducing downtimes of HVAC systems for pharmaceutical production facilities with Predictive Maintenance and Machine Learning

Niels Humbeck

32003045

27.08.2021

Cologne, Germany

1 Table of contents

2 Table & Figure directory

3 Abbreviations

Table 1: Abbreviations

Abbreviation	Meaning
HVAC	Heating, Ventilation and Air Conditioning
IoT	Internet of Things
OEE	Overall Equipment Effectiveness
OOS	Out of Specification
ROI	Return on Investment
MLC	Machine learning canvas
QA	Quality Assurance
PLC	Programmable logic controller
MES	Manufacturing execution system
ERP	Enterprise resource planning
CSF	Critical success factor
KPI	Key performance indicator

4 Motivation

Common metrics to assess the performance of a pharmaceutical production site are overall equipment effectiveness (OEE), productivity and yield. One reason for a bad performance in this metrics are frequently room condition being out of specifications (OOS). HVAC (heating, ventilation & air conditioning) is an integral part of the pharmaceutical production controlling the room conditions. To comply with GMP[1]-Guidelines, the room conditions specified in the manufacturing license must be met (e.g. see WHO 2010) . HVAC systems are fulfilling the following tasks (see Zaki 2015, Choudhary 2021): Reduction of airborne particles (dust & micro-organism), maintaining the pressure cascades to avoid contamination, regulating the room temperature, regulating the room humidity.

Maintaining the HAVAC system is a key factor to assure availability. In many sites the maintenance is still done in a reactive way (run to a failure after breakdown), or maintenance is frequently scheduled (preventive maintenance). With an increase of sensor data condition based monitoring is becoming more popular (Cinar 2020). The maintenance is done after the sensors report a deterioration of one or more parameters. With the help of machine learning models the date of a machine breakdowns can be predicted. Predicting the breakdown of a machine in order to plan maintenance activities is called predictive maintenance.

The idea of my project is to found a Startup called “HVAC-Angel” to offer solutions to increase the availability of the HVAC system with predictive maintenance and machine learning to boost OEE, productivity and reduce yield losses. Necessary information are gathered by the means of IoT sensors. The target is to detect defects before they cause breakdowns in order to plan maintenance on time.

5 The Machine Learning Canvas

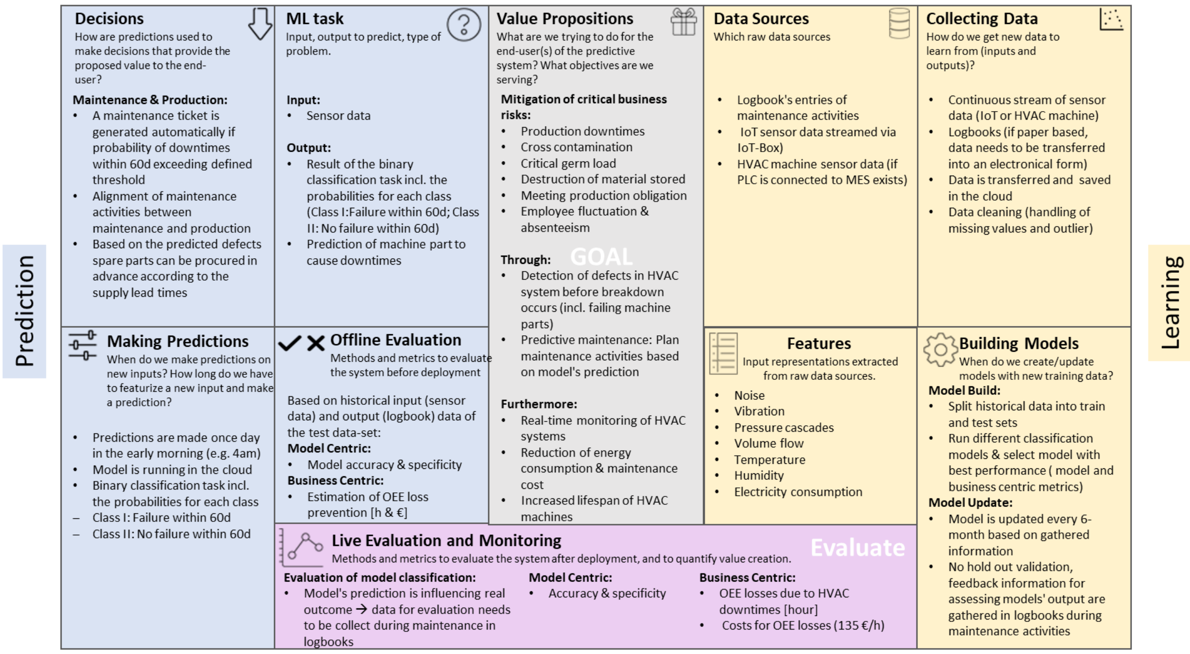
Up to 80-85% of the data science projects are never making it into production. From the 15-20% of completed data science projects it is assumed that only 8% generates values to the company and 92% not (Thomas 2020). Completing a data science projects which adds no value to the business is a waste of resources. One of the main reasons is that not the right use cases for data since projects were found (Guasduff 2019). Focusing in the very beginning of a data science project on the use case can prevent the project to fail. Not applying a use case analysis before implementation causing following typical errors:

In many cases the data scientists and the business stakeholder are disconnected and not aligned on a shared goal (Dorard 2016). The data scientists can build the best model on earth

[1] Good manufacturing practice: Requirements to produce drugs in high quality

but if its not adding a surplus value to the business by answering a critical question it is still useless and won't get deployed in production or even worse deployed in production but not used (Thomas 2020). Adopting new technologies needs a strong commitment from the business stakeholder, otherwise the model won't get used even if it adds value to the business (ODSC 2019). The value of the data science project should not be overestimated otherwise the business stakeholder getting disappointed and will not support further projects (ODSC 2019). Other risks are the scope and the quality of the data. Starting a data science project without knowing which data in which quality is available results frequently in an early project stop. In some cases the data needs to be labeled by human workforces (e.g. Appen). The cost for labeling the data exceeds frequently the budget for the data science project if not considered in the very beginning (Guasduff 2019). Other risks to data science projects are security and privacy concerns (Guasduff 2019).

A solution driven approach to detect the use cases with the best chances to be deployed in production is the machine learning canvas (MLC) from Louis Dorard. The machine learning canvas is derived from the business model canvas and focusing especially on "technical aspects of implementing a machine learning based use case in a company" (Kerzel 2020). The MLC describes in a visual appealing form complex objects and topics which needs to be discussed before a data science projects is going into the implementation phase. It can be used to communicate the vision for the machine learning system to the business stakeholder to make sure that its supporting organizational objectives and to get commitment from all stakeholder. It's a first feasibility check and key constraints of the machine learning systems can be identified. In the core of the MLC is the value proposition specifying the value the data science project will generate for the business. On the right hand side the MLC describes how the model is learning (Which raw data sources do we have?, How do we get new data to learn?, Which features have the main impact on the predictions?, How do we create and update the model?). In the left side the MLC describes how the predictions of the model are made (What is the ML task?, What decisions are made based on the prediction?, When do we make predictions?, How do we evaluate the model before deployment?). In the upper part of the MLC more background information are given and in the lower part more specific points are addressed. In the bottom of the MLC the field "Live evaluation and monitoring" describes which methods and metrics are used to evaluate the machine learning model after deployment. The machine learning canvas is depicted in Figure 1 and Appendix 7.1.



*Machine learning canvas after Dorad, L (2020)

Figure 1: Machine learning canvas - HVAC Angel

5.1 Value Proposition

The machine learning model predicts the time window when a failure of a HVAC systems will cause downtimes in order to repair the HVAC machines in advance (predictive maintenance). The prediction of failures and downtimes are important for the maintenance as well as the production department to plan and align maintenance activities. Besides the predicted time window of breakdowns the machine learning model predicts the HVAC machine part most likely to break.

This solution helps our customers in the pharmaceutical industry to mitigate substantial risks to their business arising from downtimes of HVAC systems:

- Downtimes in production since the required room condition are not met. The OEE as well as the productivity are going to decrease significantly
- Increased yield losses (e.g. packaging material is deformed)
- If the storage conditions are not met, stored material need to be destructed resulting in a financial loss
- The ability of the production site to meet the production obligations are going to be questioned if frequent downtimes occur
- Germ growth in hidden places is difficult to identify and to clean resulting in worst case in a production shutdown
- If pressure cascades are insufficient cross-contamination can occur

- Employees wellbeing is negatively impacted resulting in an increased absenteeism and fluctuation rate.
- Up to 50% of the energy consumption of a production site is consumed in the HVAC systems (Zaki 2015). A defect HVAC system will consume considerable more energy

The main customer target group are pharmaceutical production sites with older HVAC systems with no or not connected sensor data. A second industry segment are the electronic components manufacture, since the requirements for room conditions are similar. In the mid-term a collaboration with HVAC equipment suppliers to get the solution installed in new HVAC systems has to be assessed.

The offered solution includes the sale of IoT-devices to gather the relevant information, the installation services as well as the software to predict maintenance. The IoT devices and installation expenses are onetime costs. For the software a license model is offered. The amortization time for our customer is 10 month and the ROI over 5 years is 285% for one HVAC system. The profitability calculation includes only the benefits of saving salaries and line amortization due to reduced downtimes. If other factors (e.g. lost sales, indirect labor costs, reduced maintenance costs and increased lifespan of HVAC machines) are included the ROI will be significantly higher. HVAC-Angel will generate 21k€ in sales for one HVAC system over 5 years (see Appendix 1.1).

Alternatives to the offered solution are:

- Keep the existing maintenance mode, which results in financial losses as outlined in Appendix 1.1.
- Increase the frequency of scheduled (preventive) maintenance, which results in higher cost for maintenance as well as in downtimes during maintenance.

In comparison to this alternatives the offered solution is competitive. Furthermore the solution enables real-time monitoring and evaluation of room conditions. It can be integrated – if QA-approved - into the electronical batch records to automate quality assurance as well as production processes. The life time of HVAC systems can be increased with predictive maintenance by 20 to 40 percent (Dilda 2017) and the cost for maintenance can be reduced as well by 25 to 30 percent (Poor 2019).

5.2 Data Sources

There are different raw data sources to fed the machine learning algorithm. First of all real-time data of IoT sensors are fed into the machine learning algorithm. The latency of this data is low (<0.5 s). Depending on the sensor type the volumes and velocity can fluctuate. Another source of data can be the HVAC machine data if a PLC connected to a MES system is in place. This data is similar to the IoT sensor data. These sensor data are coming along with a time stamp. A valuable source for the condition of the HVAC machines are logbooks. For pharmaceutical production it is strictly required to log all information of maintenance activities as well as downtimes events of HVAC systems in a logbook. The status of maintenance checkups and or breakdowns/ failures are listed including a times stamp (e.g. breakdown component and measurement to fix it). The logbook data might be analog (hand written text book) and needs to be transferred into a digital file. It is recommend to collect information about the maintenance in a digital form in order to avoid translation costs after deployment.

5.3 Collecting Data

Input data is collected from HVAC machine- or IoT- sensor data in order to build the initial proof of concept model. The IoT sensor data is uploaded via a IoT-box into a cloud. The HVAC machine sensor data are either collected via a USB-download or via a PLC connected to a MES system. The output data are included in the logbooks (e.g. explicit failure description with timestamp). As in 5.2 mentioned in most cases the logbook data needs to be transferred into a digital file.

After deployment the IoT sensors stream the data via IoT-boxes into the cloud. In some cases data from machine sensor can be used when a PLC connected to a MES exist. In those cases no IoT sensors needs to be installed. The machine sensor data is collected in the cloud as well. Output data are gathered in the logbook. It is recommend to use an electronic logbook. The logbook contains all information about every failures and downtimes as well as information about the result of predicted maintenance activities. It is important to access whether predictive maintenance activity were necessary. The machine learning model predicts the failure chance of a specific machine part. After the maintenance activities the plausibility of the alert needs to be assessed: What was the status of the machine part predicted to fail? Was the alert reasonable or were no failure visible? Those output data needs to be collected in order to access the quality of the prediction model and to improve the machine learning model by retraining.

The input data is “automatically cleaned, organized, standardized and checked for errors” (Kerzel 2020B). Especially a focus on missing values as well as outliers is set.

5.4 Features

The machine learning model use following features to predict failures and downtimes:

- Noise [numerical]

- Vibration [numerical]
- Pressure [numerical]
- Volume flow [numerical]
- Temperature [numerical]
- Humidity [numerical]
- Electricity consumption [numerical]

All features are time series and collected either from the IoT sensor or the HVAC machine sensors (via PLC connected to MES).

5.5 Building Models

In literature two main approaches for building a model to predict breakdowns in order to enable predictive maintenance are discussed.

1) Phenomenological (or data-driven) approach.
   A failure is detected "by learning the difference between normal and anomalous behaviors from the historical data" (Satta n.a).
2) Model-based approach.
   A simulation model of the existing HVAC system is generated (e.g. based on schematics and blueprints) and a "failure is detected by comparing the actual behavior with the simulated one" (Satta n.a).

The first phenomenological approach is chosen science the information of the HVAC system to create a simulation model are in most cases not available.

In order to build a machine learning model to predict failures and breakdowns different machine learning algorithms (classification task) will be trained based on the gathered historical input data (see 5.3). The logbook data are used as output data. The different models will be evaluated by the measures (model centric as well as business centric) described in 5.9 and the model with the highest scores will be chosen.

The models needs to be retrained frequently. A first estimation of the frequency of retraining are every 6 month. There are two reasons for retraining the model. First we collect over the time more data and with more data we can improve our prediction model. And second the dynamics can change and the model needs to be adjusted in order to prevent a decay in quality (Dorard 2017). For the retraining/ updating of the model are the digital logbook data from the maintenance activities are particular important as described in 5.3.

5.6 ML-Task

As described in 5.2 and 5.3 the sensor data are streamed in near real-time via a IoT-box or a PLC connected with MES into the cloud. The machine learning model uses the sensor data as an input for predictions. The machine learning algorithm solves a binary classification task.

- Class 1: Breakdown occurs in the next 60 days
- Class 2: Breakdown will not occur in the next 60 days

In further development more classes can be distinguished. The model predicts based on the input data the likelihood for class 1 or class 2. If class 1 is more likely and exceeds a defined threshold the machine learning model predicts the HVAC machine part most likely to cause the breakdown.

As a baseline historical data can be applied like MTBF (Mean time between failure).

5.7 Decision

If the prediction probability of Class 1 (see 5.6) exceeds a defined threshold (depending on the criticality of the production lines) a maintenance ticket is generated automatically. It is recommended to implement an interface to the customers maintenance ERP module. In the maintenance ERP module or the online effort dashboard the maintenance ticket is visible for the maintenance department as well as to the production department. If no interface towards the ERP system exists the ticket will be shown in an online dashboard. A E-Mail notification can be implemented. The maintenance ticket includes information about the machine part predicted to cause downtimes, the time window of the predicted downtimes (Class 1) and the probability of this event. Based in this maintenance ticket the exact maintenance date can be aligned between the maintenance department and the production department. Based on the ticket spare parts will be ordered by the maintenance department, if necessary. The final decision is not made by the machine learning model but by the maintenance department. To increase the trust in the predictions a probability threshold is defined as described above as well as an anomaly score of each prediction. If it is a rare event not well covered by the training data (high anomaly score) the prediction has a higher uncertainty.

No hold out data set are applied in this use case since the HVAC systems are business critical and feedback to improve the model is generated in the logbooks (see 5.10).

5.8 Making Predictions

The predictions are generated once a day. Preferable in the early morning (e.g. 4 am) in order to provide the maintenance and the production department with updates before the working day starts. The machine learning model is running on cloud resources since the amount of data is too large for an edge solution. The input data are stored in the cloud as well. Cloud solutions increase the horizontal scalability. The prediction tasks for all supervised HVAC systems are expected to be processed within 5 - 30 minutes. The results in from of the maintenance ticket are shown in an online dashboards or in the maintenance ERP module (if an interface to the customers ERP solution exist). Every customer has only access to the prediction of their HVAC system. The maintenance tickets includes information about the machine part predicted to cause downtimes, the time window (see 5.6) of the predicted failure and the probability of this event

It is possible to create alerts in the real-time monitoring system. If a parameter drops or exceeds a specified parameter the maintenance department is informed (not a machine learning prediction).

5.9 Offline Evaluation

Before the model is deployed it is tested extensively. As in 5.2 and 5.3 explained the input data are gathered either via the IoT sensors or via the machine sensors. The Output data is gathered from the logbooks of the HVAC machines. The data is split into a train and test data set. The model is trained with the training data set and evaluated with the test data set. The model centric evaluation methods is described in 5.10. It is recommend to check the business centric evaluation for each customer before implementation (see 5.10). After the accuracy and specificity of the model exceed a defined threshold (model centric) and the monetized outcome (business centric) exceed a defined threshold as well the model is approved to be used in production.

The model needs to be retrained every 6 month (see 5.5) and will be evaluated after each restraining again. For this retraining the sensor data as well as the logbooks entries are used to increase the accuracy.

5.10 Live Evaluation and Monitoring

There are two perspectives to evaluate machine learning models: The business centric evaluation and the model centric evaluation.

The **business centric evaluation** focusing on the wider implication of the machine learning model to business metrics and analyze how much value is generated for the business by using the predictions (Kerzel 2020). As in 5.1 explained the predictions of the machine learning model mitigates business critical risks. Mitigating these risks are critical success factors (CSF) for every pharmaceutical production site assuring the competitiveness and future viability. Key performance indicators (KPI) translate the CSF into measurable quantities (Kerzel 2020). The impact of the described risks on the business like downtime in production, ability to meet production obligations, production stops due to germ contamination, employees fluctuation and absenteeism can be measured through the overall equipment effectiveness (OEE), which is already a wide-spread indicator for production sites. It is common to break down the OEE into the main loss reasons. The main business KPI for evaluating the impact of the machine learning models predictions is the OEE-losses due to HVAC failures. The target is to reduce the OEE losses in hour[2] by 50% in comparison to the average losses in the last 2 years (see McKinsey study: Dilda 2017). The ROI of the data science project can be estimated by monetarizing the reduction in OEE losses. As in 1.1 explained one hour of OEE losses causes costs of 135€ just for amortization and personal costs exclusive lost sales, costs for indirect employees, increased lifespan of HVAC systems nor the reduction of maintenance costs. Thus the business calculation is conservative. The business benefit is calculated as the product of OEE losses [hour] prevented and the cost for one OEE loss hour. The OEE losses caused by HVAC failures are one of the main KPI for the maintenance department to be tracked every day. The production is discussing this KPI in the daily performance meeting on line-level (for each line) and production unit level (for all lines) as well. On site level the overall OEE is discussed every day.

The **model centric evaluation** is focusing on the question: Are the predictions of the machine learning model reliable and accurate (Kerzel 2020). In the first development phase the machine learning model is solving a binary classification task with the following possible outcomes (see Binding 2019):

- True positive (TP): Failure exist and model detected a failure - Maintenance is scheduled (1-2h of planned downtimes); unplanned downtimes are prevented [high benefit for business]
- False positive (FP): No failure exist but model detected a failure - Maintenance is scheduled (1-2h of planned downtimes); no unplanned downtimes occurs [low cost for business]
- True negative (TN): Failure does not exist and model does not detect a failure- Maintenance is not scheduled; no planned or unplanned downtimes occurs [no costs and no benefits for business]
- False negative (FN): Failure exists but model did not detect a failure - Maintenance is not scheduled; unplanned downtimes occur [high costs for business]

[2] OEE is the ration of useful time divided by the net opening time. The difference between net opening time and the useful time are the OEE losses in hour. These losses are broken down in categories (e.g. downtimes due to room condition).

The performance of the classification model is measured by two measures: Accuracy and Specificity. The overall correctness of the model is measured by the $\text{accuracy} = \frac{TP + TN}{TP+TN+FP+FN}$. As described above the false negatives will lead to a large negative business impact. Therefore the model has to be trained to avoid false negative predictions. Therefore a key measurement is $specificity = \frac{TN}{TN+FN}$. A high specificity indicates that a negative prediction of the model is classified correctly (TN) in relation to all negative predictions.

An evaluation and retraining of the model is done every 6 month. The predictions of the machine learning model are influencing the outcome, as the predictive maintenance prevents failures to result in downtimes. Hold-out validation is not applicable in this use case since failures are a rather rare event with a possible high negative impact to business metrics (e.g. OEE losses). In order to evaluate the performance of the machine learning predictions the maintenance team has to provide further information during the maintenance activities: Was the maintenance necessary? Was the predicted time window till breakdown realistic? Was the identified machine part responsible for the failure?

6 High level technical concept

The high level technical concept is depicted in Figure 2.

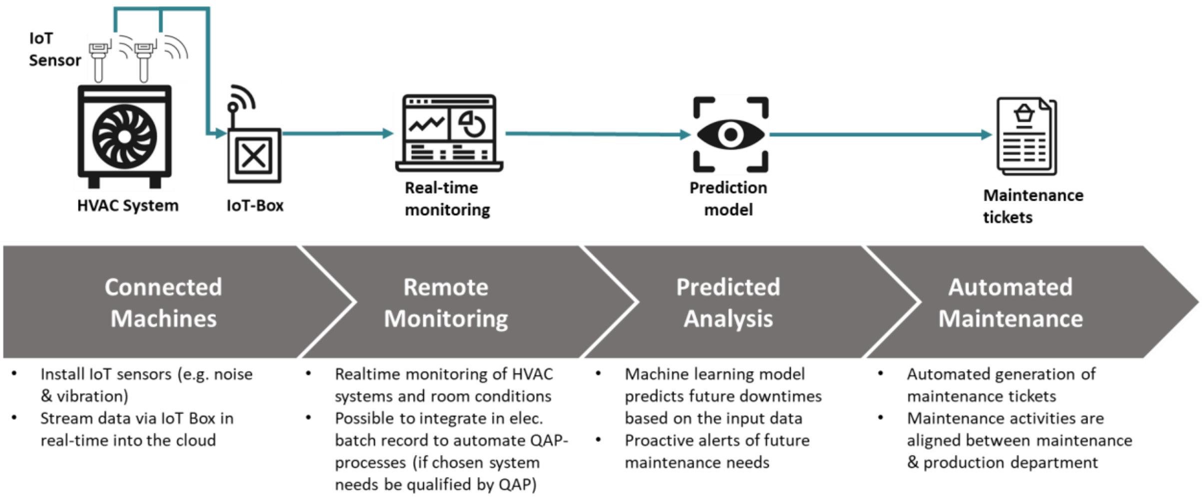
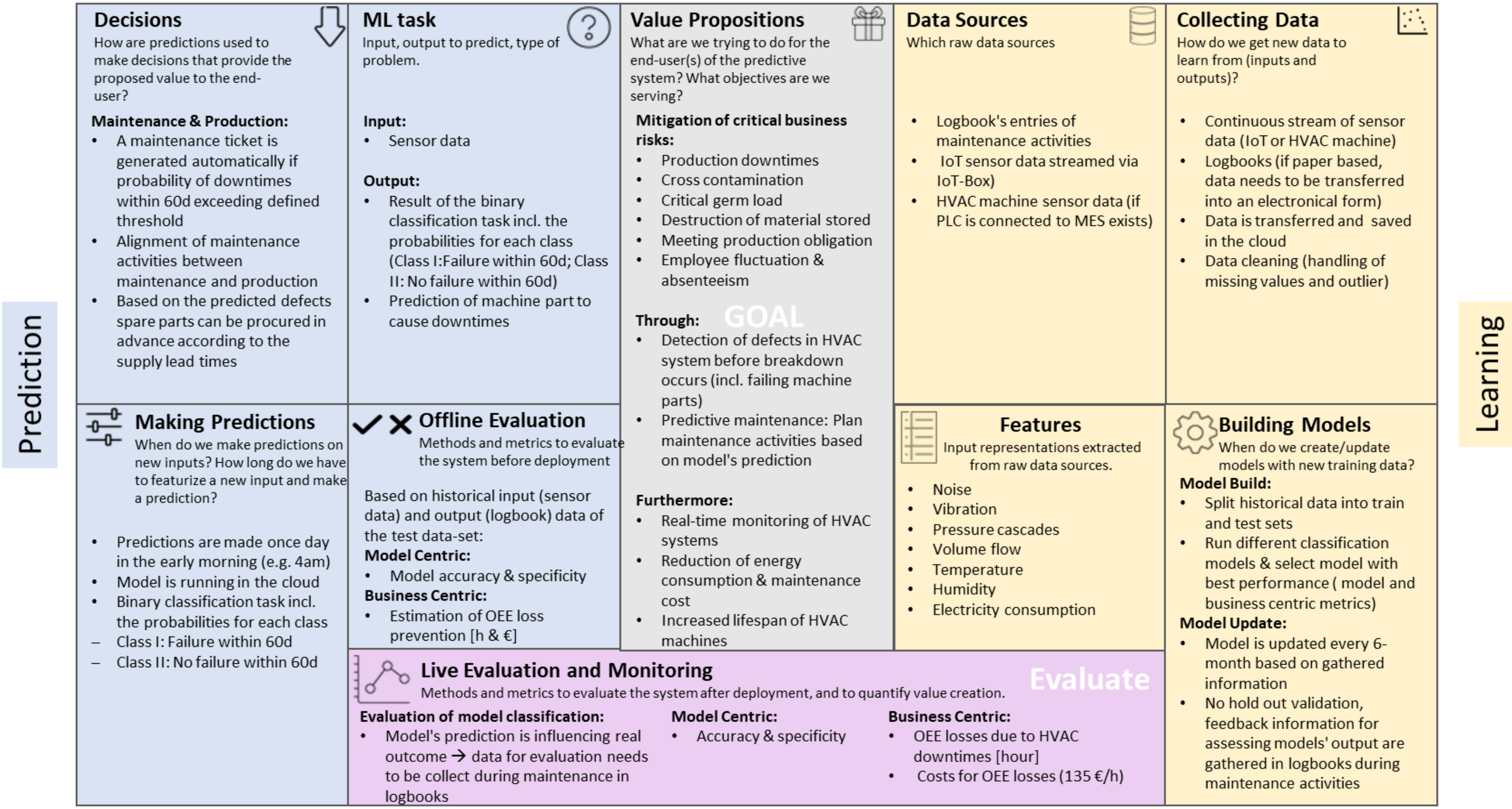


Figure 2: High level technical concept HVAC-Angel

IoT sensors (e.g. noise, vibration, temperature etc.) are installed in the HVAC machines. The IoT sensor data is streamed in real-time via a IoT-Box into the cloud to be saved. If machine sensors exist which have PLC and MES connection, these data can be transferred into the cloud as well. The sensor data are time series which means for every data point exist a timestamp. The sensor data are visualized in a remote real-time monitoring dashboard. It is possible to create alerts if parameters fall below or exceed a defined threshold. The real-time monitoring system can be integrated into an electronic batch record system in order to automate quality assurance processes. A prerequisite is the qualification of the real-time monitoring system by quality assurance to use this data for the electronic batch record. Once a day the prediction are made in the cloud based on the saved IoT sensor data. Depending on the amount of data and HVAC systems the computational resources can be scaled easily. If the machine learning model predict a failure causing downtimes in the future, a maintenance ticket is created (see 5.8). This maintenance ticket can be depicted in an online dashboard or in fed in the ERP maintenance module if a interface exists. Maintenance activities can be aligned between maintenance department and production department, based on the generated maintenance ticket.

7 Appendix

7.1 Machine learning Canvas



*Machine learning canvas after Dorad, L (2020)

7.2 Return on investment calculation

Assumption:

- 1 HVAC system for 2 (filling &) packaging lines
- 3 line workers for 1 (filling &) packaging lines
- 3 shifts per day
- Salary of a line employees (E7 4 years see Böckler 2019): 3600€ gross salary + 900 € employer contribution: 4500 €/month ~30 €/hour
- Amortization of a (filling &) packaging lines: 6 Mio. € in 15 years → 1.200 €/day
- 5 days a year unplanned downtimes due to OOS room conditions
- 50% of unplanned downtimes can be avoided with the help of the offered solution

Costs for predictive maintenance solution for 1 HVAC system:

- IoT devices: 6.000 € (inclusive margin)
- Software license: 2000 €/year
- Average installation cost: 5.000 €
- ➢ Cost in the first year: 13.000 €
- ➢ Costs in the following years: 2.000 €/year

Not included:

- Loss of sales due to reduction in production volumes
- Costs for indirect employees
- Increased time for finished good release time due to deviations management

Cost for Salary per line and day:

$$C_{SalaryPerLineAndDay} = 3\,\frac{shifts}{day} * 3\,\frac{employees}{line} * 7{,}5\,\frac{h}{shift} * 30\,\frac{Euro}{h} = 2.025\,\frac{Euro}{line\ \&\ day}$$

Cost for amortization per line and day:

$$C_{AmortizationPerLineAndDay} = 1.200\,\frac{Euro}{line\ \&\ day}$$

Cost savings due to predictive maintenance solution per year:

$$\frac{Savings}{year} = 50\% * 2\ lines * 5\ days * \left[\ C_{SalaryPerLineAndDay} + C_{AmortizationPerLineAndDay}\right]$$

$$\frac{Savings}{year} = 50\% * 2\ lines * 5\ days * \left[2.025\,\frac{Euro}{line\ \&\ day} + 1.200\,\frac{Euro}{line\ \&\ day}\right]0 = 16.125\ Euro$$

Amortization of Invest:

$$T_{Amortization} = \frac{\text{Cost in the first year}}{\frac{Savings}{Year}} = \frac{13.000\ Euro}{16.125\ Euro} * 12\ Month = ca.10\ Month$$

Return on Invest in 5 years:

Table 2: ROI - Costs, Savings and Saldo

Year	1	2	3	4	5	Sum
Costs	13.000 €	2.000 €	2.000 €	2.000 €	2.000 €	21.000 €
Savings	16.125 €	16.125 €	16.125 €	16.125 €	16.125 €	80.625 €
Saldo	3.125 €	14.125 €	14.125 €	14.125 €	14.125 €	59.925 €

$$ROI = \frac{\sum Savings - \sum Costs}{\sum Costs} = \frac{59.925\ Euro}{21.000\ Euro} = 285\%$$

(no inflation or interest rates considered)

8 Literature

- Bhati, A. (n.a.), HVAC Design for Pharmaceutical Facilitis, CED engineering
- Binding, A., Dykeman, N., Pang, S., (2019), Machine Learning Predictive Maintenance on Data in the Wild, 2019 IEEE 5th World Forum on Internet of Things (WF-IoT)
- Böckler (2019), Trarifentgelte Chemische Industrie, https://www.boeckler.de/pdf/p_ta_tarife_chemische_industrie_2019.pdf, last access 22.05.2021 at 1pm
- Chin, J., lin, S., Persada, S., Jaqin, C., Mufidah, I. (2020), Preventive maintenance model for heating ventilation air conditioning in pharmacy manufacturing sector, The Society for Reliability Engineering, Quality and Operations Management (SREQOM), India
- Choudhary, A. (2021), Working of HVAC System and Its Functions, https://www.pharmaguideline.com/2017/09/working-of-hvac-system-and-its-functions.html, last access 22.05.2021
- Cinar, Z., Nuhu, A., Zeeshan, Q., Korhan, O., Asmael, M., Safaei, B. (2020), Machine Learning in Predictive Maintenance towards Sustainable Smart Manufacturing in Industry 4.0
- Dilda, V., Mori, L., Noterdaeme, O., Schmitz, C. (2017), “Manufacturing: Analytics unleashes productivity and profitability”, https://www.mckinsey.com/business-functions/operations/our-insights/manufacturing-analytics-unleashes-productivity-and-profitability#, last access 02.06.2021 at 23pm
- Dorad, L (2020): Machine Learning Canvas (https://www.louisdorard.com/machine-learning-canvas) [last accessed: 03.06.2021]
- Dorard, L. (2016), From Data to AI with the Machine Learning Canvas (Part I)- A framework to connect the dots between data collection, machine learning, and value creation, https://medium.com/louis-dorard/from-data-to-ai-with-the-machine-learning-canvas-part-i-d171b867b047, last access 27.05.2021 at 20pm
- Dorard, L. (2017), From data to AI with the Machine Learning Canvas (Part II), https://www.linkedin.com/pulse/from-data-ai-machine-learning-canvas-part-ii-louis-dorard/, last access 27.05.2021 at 21pm
- Dorard, L. (2018A), From Data to AI with the Machine Learning Canvas (Part III), https://www.linkedin.com/pulse/from-data-ai-machine-learning-canvas-partiii-louis-dorard/ , last access 27.05.2021 at 21pm

- Dorard, L. (2018B), From Data to AI with the Machine Learning Canvas (Part IV), https://www.linkedin.com/pulse/from-data-ai-machine-learning-canvas-part-iv-louis-dorard/ , last access 27.05.2021 at 21pm
- Guasduff, L. (2019), 3 Barriers to AI [/ML] Adoption, https://www.gartner.com/smarterwithgartner/3-barriers-to-ai-adoption/, last access 29.05.2021 at 6:26pm
- Guo, Y., Wall, J., Li, J., West, S. (2013), Intelligent Model Based Fault Detection and Diagnosis for HVAC System Using Statistical Machine Learning Methods
- Icon 1: https://thenounproject.com/term/angel/634522/ , last access 22.05.2021
- Icon 2: https://thenounproject.com/search/?q=hvac&i=2249022 , last access 22.05.2021
- Kerzel U. Prof-Dr. (2020B), Data Science, IUBH Internationale Hohschule GmbH
- Kerzel, U. Prof. Dr. (2020), Use case and evaluation, IUBH Internationale Hohschule GmbH
- Liulys, K. (2019), Machine Learning Application in Predictive Maintenance, Department of Electrical Engineering Vilnius Gediminas Technical University
- ODSC (2019), 6 reasons why dat science projects fails, https://medium.com/@ODSC/6-reasons-why-data-science-projects-fail-6240bf9326f6, last access 29.05.2021 at 5 pm
- Poor, P., Zenisek, D., Basl, J. (2019), Historical Overview of Maintenance Management Strategies: Development from Breakdown Maintenance to predictive maintenance in accordance with four industrial revolutions, Proceedings of the International Conference on Industrial Engineering and Operations Management Pilsen, Czech Republic
- Satta, R., Cavallari, S., Pomponi, E., Grasselli, D., Picheo, D., Annis, C. (n.a.), A dissimilarity-based approach to predictive maintenance with application to HVAC systems
- Sitton-Candanedo, I., Rodriguez-Gonzales, S., Corchardo-Rodrigues, J. (n.a.), Fault predictive model for HVAC Systems in the context of Industry 4.0, Asociacion Espanola Para La Inteligencia Artificial
- Srinivasan, R., Islam, T., Wang, Z., Sookoor, T. (n.a.), Preventive Maintenance of Centralized HVAC Systems: Use of Acoustic Sensors, Feature Extraction, and Unsupervised Learning; Department of Computer Science, University of Houston
- Thomas (2020), 10 reasons why data science projects fails, https://fastdatascience.com/why-do-data-science-projects-fail/ , last access at 29.05.2021 at 5:18 pm

- VB (2019), Why do 87% of data science projects never make it into production?, https://venturebeat.com/2019/07/19/why-do-87-of-data-science-projects-never-make-it-into-production/, last access 29.05.2021 at 7:11 pm
- WHO (2010) , WHO good manufacturing practices for pharmaceutical products containing hazardous substances, http://academy.gmp-compliance.org/guidemgr/files/WHO_TRS_957_ANNEX3.PDF, last access 22.05.2021
- Xuemei, L., Ming, S. Lixing, D. (2009), HVAC Fault Diagnosis System Using Rough Set Theory and Support Vector Machine, Second International Workshop on Knowledge Discovery and Data Mining
- Yarmoluk, D., Truempi, C. (2019), Why move from condition monitoring to predictive maintenance?, https://www.ibm.com/blogs/internet-of-things/iot-condition-monitoring-part-one/, last access 22.05.2021 at 11am
- Zaki, M., Mishra, D. (2015), Overview of HVAC System: Operational Significance of HVAC Provision for Pharmaceutical Facilities, International Journal on Recent and Innovaion Trends in Computing and Communication

8.1 Icons

- https://thenounproject.com/term/ventilator/2589871/
- https://thenounproject.com/search/?q=IoT+Sensor&i=3940622
- https://thenounproject.com/search/?q=IoT+box&i=2738695
- https://thenounproject.com/search/?q=analysis&i=1790270
- https://thenounproject.com/term/vision/286423/
- https://thenounproject.com/term/order/2294578/
- https://thenounproject.com/term/angel/634522/
- https://thenounproject.com/search/?q=hvac&i=2249022

HVAC Angel

Pitch Deck: Reducing downtimes of HVAC systems for pharmaceutical production facilities with Predictive Maintenance and Machine Learning

Niels Humbeck (CEO)

30.07.2021 Köln

Agenda

1. Problem description
2. Our solution
3. Our promise
4. Our service
5. How can you contribute?

Old HVAC-Systems can lead to a multitude of risks…

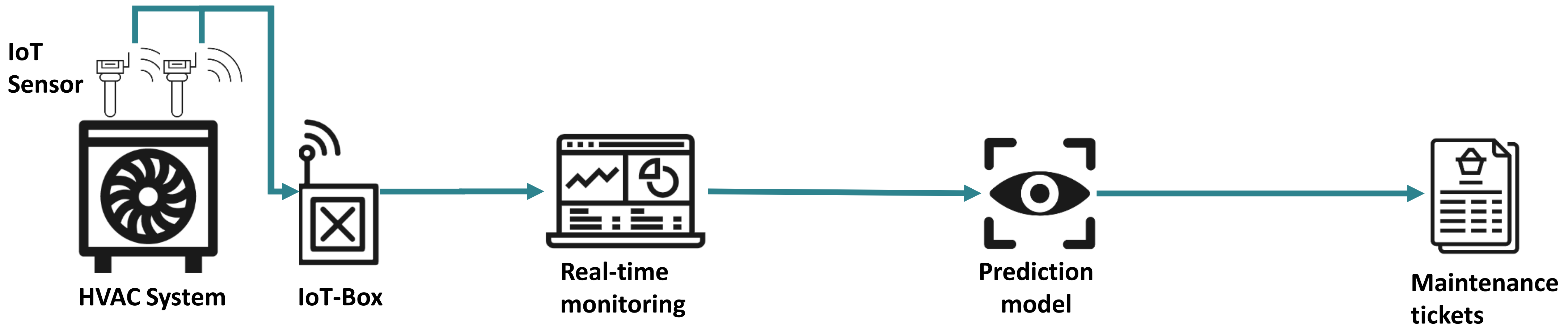
Breakdown of HVAC system is business critical for pharmaceutical companies

HVAC Breakdowns

- Losses in sales
- Decrease in OEE
- Increased yield losses
- Cross-contamination
- Destruction of goods
- Quality violations
- High energy consumption
- Employee fluctuation

Our solution

Predicting HVAC breakdowns be the mean of IOT devices and machine learning



Connected Machines	Remote Monitoring	Predicted Analysis	Automated Maintenance
• Install IoT sensors (e.g. noise & vibration) • Stream data via IoT Box in real-time into the cloud	• Realtime monitoring of HVAC systems and room conditions • Possible to integrate in elec. batch record to automate QAP-processes (if chosen system needs be qualified by QAP)	• Machine learning model predicts future downtimes based on the input data • Proactive alerts of future maintenance needs	• Automated generation of maintenance tickets • Maintenance activities are aligned between maintenance & production department

See Cinar (2020)

Our promise

Reduction of unplanned downtimes due to HVAC breakdowns by up to 50%

~ 50% reduction in downtimes (HVAC)

See ROI calculation in back up

ROI + 285% over 5 years

See ROI calculation in back up

~ 10-month amortization

See ROI calculation in back up

HVAC Angel

Real-time monitoring

Increased service time by 20-40%

See (Dilda 2017)

Maintenance cost reduction by up to 30%

See (Poor 2019)

Production & Maintenance partnership

Our Service

One selected supplier – everything from a single source (Sensor to Prediction)

Infrastructure provisioning

- **Supply of relevant IOT sensors**
- Installation services
- Maintenance service with low lead times

- **Pricing model:**
- One-time costs

Software Solution

- **Providing cloud-based software solutions**
- Breakdown prediction
- Alerting functionalities
- Dashboards for real-time monitoring
- Individual updated machine learning model

- **Pricing model:**
- License model (min. term 1 year)

How can you contribute?

Invest in HVAC-Angel!

Be part of our journey …

Back Up | Machine learning canvas

Prediction

Decisions
How are predictions used to make decisions that provide the proposed value to the end-user?

Maintenance & Production:
- A maintenance ticket is generated automatically if probability of downtimes within 60d exceeding defined threshold
- Alignment of maintenance activities between maintenance and production
- Based on the predicted defects spare parts can be procured in advance according to the supply lead times

ML task
Input, output to predict, type of problem.

Input:
- Sensor data

Output:
- Result of the binary classification task incl. the probabilities for each class (Class I:Failure within 60d; Class II: No failure within 60d)
- Prediction of machine part to cause downtimes

Value Propositions
What are we trying to do for the end-user(s) of the predictive system? What objectives are we serving?

Mitigation of critical business risks:
- Production downtimes
- Cross contamination
- Critical germ load
- Destruction of material stored
- Meeting production obligation
- Employee fluctuation & absenteeism

GOAL

Through:
- Detection of defects in HVAC system before breakdown occurs (incl. failing machine parts)
- Predictive maintenance: Plan maintenance activities based on model's prediction

Furthermore:
- Real-time monitoring of HVAC systems
- Reduction of energy consumption & maintenance cost
- Increased lifespan of HVAC machines

Data Sources
Which raw data sources

- Logbook's entries of maintenance activities
- IoT sensor data streamed via IoT-Box)
- HVAC machine sensor data (if PLC is connected to MES exists)

Collecting Data
How do we get new data to learn from (inputs and outputs)?

- Continuous stream of sensor data (IoT or HVAC machine)
- Logbooks (if paper based, data needs to be transferred into an electronical form)
- Data is transferred and saved in the cloud
- Data cleaning (handling of missing values and outlier)

Learning

Making Predictions
When do we make predictions on new inputs? How long do we have to featurize a new input and make a prediction?

- Predictions are made once day in the early morning (e.g. 4am)
- Model is running in the cloud
- Binary classification task incl. the probabilities for each class
  - Class I: Failure within 60d
  - Class II: No failure within 60d

Offline Evaluation
Methods and metrics to evaluate the system before deployment

Based on historical input (sensor data) and output (logbook) data of the test data-set:

Model Centric:
- Model accuracy & specificity

Business Centric:
- Estimation of OEE loss prevention [h & €]

Features
Input representations extracted from raw data sources.

- Noise
- Vibration
- Pressure cascades
- Volume flow
- Temperature
- Humidity
- Electricity consumption

Building Models
When do we create/update models with new training data?

Model Build:
- Split historical data into train and test sets
- Run different classification models & select model with best performance (model and business centric metrics)

Model Update:
- Model is updated every 6-month based on gathered information
- No hold out validation, feedback information for assessing models' output are gathered in logbooks during maintenance activities

Live Evaluation and Monitoring
Methods and metrics to evaluate the system after deployment, and to quantify value creation.

Evaluate

Evaluation of model classification:
- Model's prediction is influencing real outcome → data for evaluation needs to be collect during maintenance in logbooks

Model Centric:
- Accuracy & specificity

Business Centric:
- OEE losses due to HVAC downtimes [hour]
- Costs for OEE losses (135 €/h)

*Machine learning canvas after Dorad, L (2020)

iubh

Back Up | ROI Calculation – assumptions (1/2)

➢ **Assumption:**

- 1 HVAC system for 2 (filling &) packaging lines
- 3 line workers for 1 (filling &) packaging lines
- 3 shifts per day
- Salary of a line employees (E7 4 years see Böckler 2019): 3600€ gross salary + 900 € employer contribution: 4500 €/month ~30 €/hour
- Amortization of a (filling &) packaging lines: 6 Mio. € in 15 years → 1.200 €/day
- 5 days a year unplanned downtimes due to OOS room conditions
- 50% of unplanned downtimes can be avoided with the help of the offered solution

➢ **Costs for predictive maintenance solution for 1 HVA system:**

- IoT devices: 6.000 € (inclusive margin)
- Software license: 2000 €/year
- Average installation cost: 5.000 €
    - ➢ Cost in the first year: 13.000 €
    - ➢ Costs in the following years: 2.000 €/year

➢ **Not included:**

- Loss of sales due to reduction in production volumes
- Costs for indirect employees
- Increased time for finished good release time due to deviations management

Back Up | ROI Calculation (2/2)

Cost for Salary per line and day:

- $C_{SalaryPerLineAndDay} = 3\frac{shifts}{day} * 3\frac{employees}{line} * 7{,}5\frac{h}{shift} * 30\frac{Euro}{h} =$

  $2.025\frac{Euro}{line\ \&\ day}$

- Cost for amortization per line and day:

- $C_{AmortizationPerLineAndDay} = 1.200\frac{Euro}{line\ \&\ day}$

Cost savings due to predictive maintenance solution per year:

- $\frac{Savings}{year} = 50\% * 2\ lines * 5\ days *$

  $\left[C_{SalaryPerLineAndDay} + C_{AmortizationPerLineAndDay}\right]$

- $\frac{Savings}{year} = 50\% * 2\ lines * 5\ days * \left[2.025\frac{Euro}{line\ \&\ day} + 1.200\frac{Euro}{line\ \&\ day}\right]0 =$

  $16.125\ Euro$

Amortization of Invest:

- $T_{Amortization} = \frac{\text{Cost in the first year}}{\frac{Savings}{Year}} = \frac{13.000\ Euro}{16.125\ Euro} * 12\ Month = ca.10\ Month$

Return on Invest in 5 years:

Year	1	2	3	4	5	Sum
Costs	13.000 €	2.000 €	2.000 €	2.000 €	2.000 €	21.000 €
Savings	16.125 €	16.125 €	16.125 €	16.125 €	16.125 €	80.625 €
Saldo	3.125 €	14.125 €	14.125 €	14.125 €	14.125 €	59.925 €

- $ROI = \frac{\sum Savings - \sum Costs}{\sum Costs} = \frac{59.925\ Euro}{21.000\ Euro} = 285\%$

- (no inflation or interest rates considered)

Library (1/2)

- Dilda, V., Mori, L., Noterdaeme, O., Schmitz, C. (2017), “Manufacturing: Analytics unleashes productivity and profitability”, https://www.mckinsey.com/business-functions/operations/our-insights/manufacturing-analytics-unleashes-productivity-and-profitability#, last access 02.06.2021 at 23pm
- Böckler (2019), Trarifentgelte Chemische Industrie, https://www.boeckler.de/pdf/p_ta_tarife_chemische_industrie_2019.pdf, last access 22.05.2021 at 1pm
- https://mind-foundation.org/wp-content/uploads/2020/09/footsteps_facebook-size.jpg, image foot steps, last access 30.07.2021

Library (2/2)

https://thenounproject.com/search/?q=risk&i=3779569

https://thenounproject.com/search/?q=bin&i=4101154

https://thenounproject.com/search/?q=failed&i=3344745

https://thenounproject.com/search/?q=energy+cost&i=3302607

https://thenounproject.com/term/exit/3125650/

https://thenounproject.com/search/?q=contamination&i=4011677

https://thenounproject.com/search/?q=revenue+losses&i=2064120

https://thenounproject.com/search/?q=throw+away&i=3168023

https://thenounproject.com/search/?q=slower&i=4009774

https://thenounproject.com/term/roi/2514388/

https://thenounproject.com/term/roi-dollar/1339611/

https://thenounproject.com/term/battery/112584/

https://thenounproject.com/search/?q=boost&i=337973

https://thenounproject.com/term/shield/1630365/

https://thenounproject.com/term/cost-reduction/3189287/

https://thenounproject.com/search/?q=Partnership&i=4079214