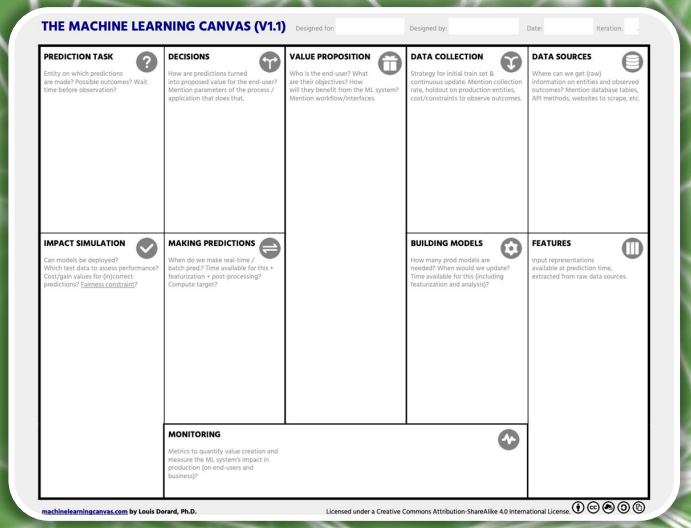
Das Machine Learning Canvas

Systematische Erhebung notwendiger Informationen zur Bewertung von Machbarkeit und Potenzial von KI Use Cases

Autorin:
Monika Arbter-Hubrich, Stuttgarter WiMa Tage 2024



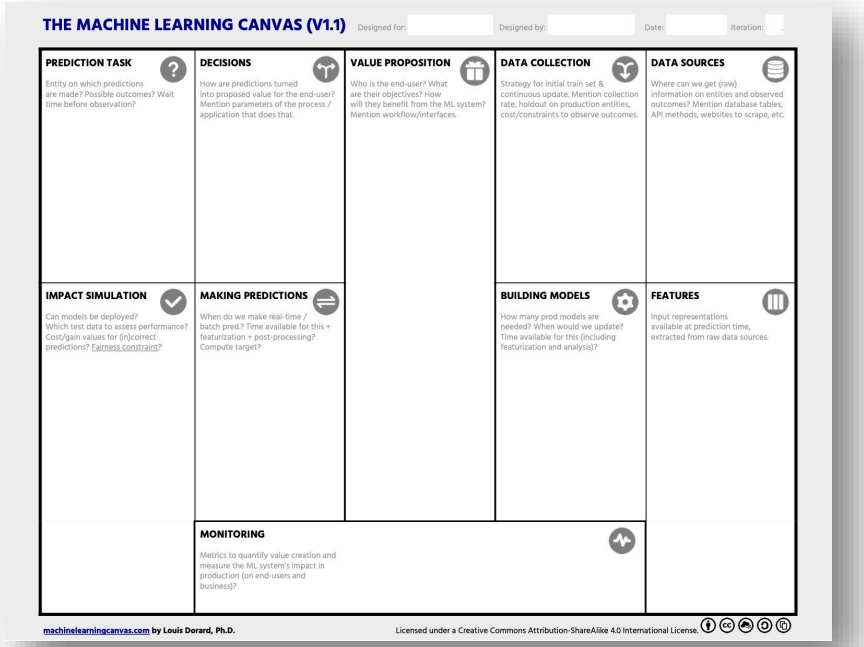
Über mich

- Monika Arbter-Hubrich
- Geboren 1983
- Meine Leidenschaft
  Digitalisierung im Kundenservice

Wissenschaftliche Arbeit trifft auf praktische Relevanz im Unternehmen

1. Überblick

- Entwickelt von Louis Dorard Ph.D.
- Angelehnt an das Business Model Canvas
- Strukturelles Werkzeug, um Machine Learning Projekte effektiv zu planen und umzusetzen



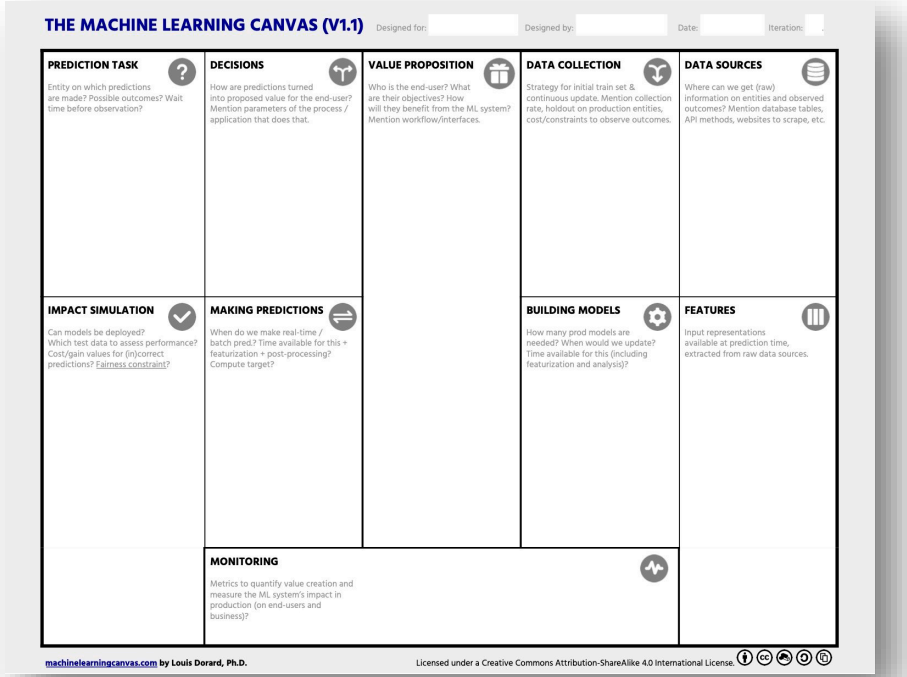
The Machine Learning Canvas, 2015
https://www.ownml.co/machine-learning-canvas

Used and recommended by thousands of teams worldwide

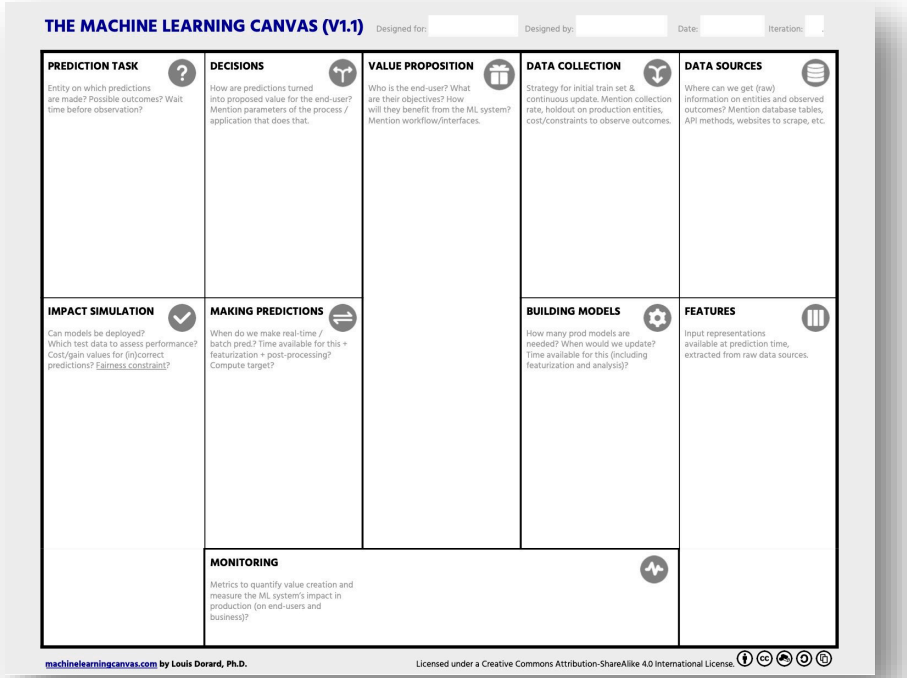
aws ESADE Business School HITACHI UCL Deutsche Telekom SAMHAMMER The Service Excellence Company

2. Zielsetzung

- Bedarf an klarer Struktur und Kommunikation in ML-Projekten
- Unterstützung interdisziplinärer Teams (Data Scientists, ML Engineers, Softwareentwickler, Projektmanager, Stakeholder,…)
- Übergang von Potenzialbewertung zur Umsetzung ohne Informationsverluste



3. Aufbau



THE MACHINE LEARNING CANVAS (V1.1) Designed for: Designed by: Date: Iteration:

PREDICTION TASK
Entity on which predictions are made? Possible outcomes? Wait time before observation?

DECISIONS
How are predictions turned into proposed value for the end-user? Mention parameters of the process / application that does that.

VALUE PROPOSITION
Who is the end-user? What are their objectives? How will they benefit from the ML system? Mention workflow/interfaces.

DATA COLLECTION
Strategy for initial train set & continuous update. Mention collection rate, holdout on production entities, cost/constraints to observe outcomes.

DATA SOURCES
Where can we get (raw) information on entities and observed outcomes? Mention database tables, API methods, websites to scrape, etc.

IMPACT SIMULATION
Can models be deployed? Which test data to assess performance? Cost/gain values for (in)correct predictions? Fairness constraint?

MAKING PREDICTIONS
When do we make real-time / batch pred.? Time available for this + featurization + post-processing? Compute target?

BUILDING MODELS
How many prod models are needed? When would we update? Time available for this (including featurization and analysis)?

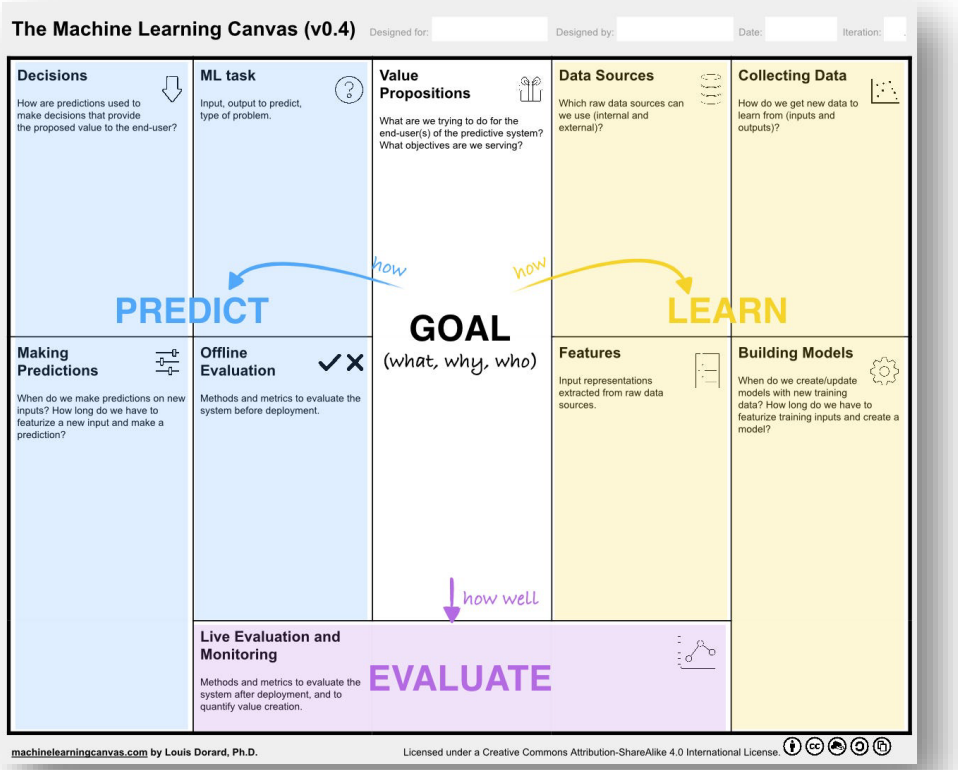
FEATURES
Input representations available at prediction time, extracted from raw data sources.

MONITORING
Metrics to quantify value creation and measure the ML system's impact in production (on end-users and business)?

machinelearningcanvas.com by Louis Dorard, Ph.D.

3. Aufbau

- Kernbereiche
  I. Goal (what, why, who)
  II. Learn (how)
  III. Predict (how)
  IV. Evaluate (how well)



3.1 Value Proposition (what, why, who)

Hauptfragen zur Value Proposition

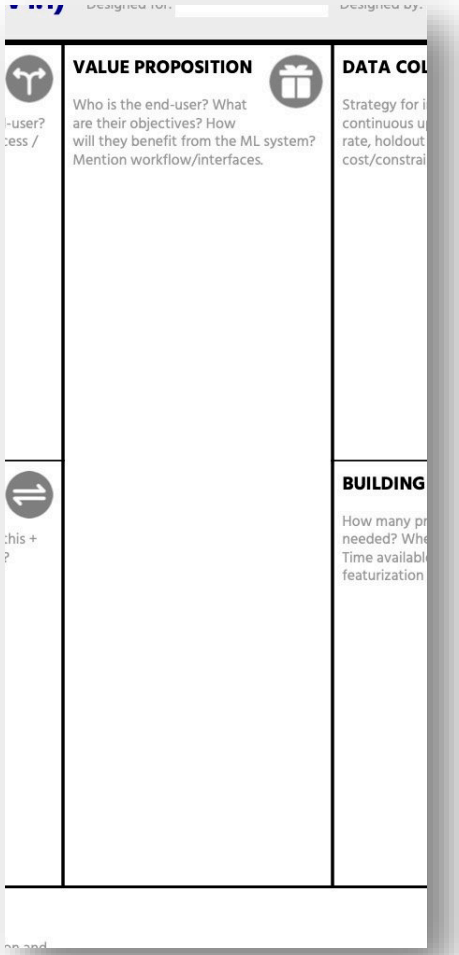
- Welches Problem wird durch das ML-Projekt gelöst?
  Detaillierte Beschreibung der Aufgabenstellung
- Welche spezifischen Vorteile bietet die Lösung den Nutzern und wer sind die Nutzer?
- Welche spezifischen Vorteile bietet die Lösung dem Unternehmen?
- Welche messbaren Ergebnisse oder Verbesserungen sind zu erwarten?
  (z.B. Kostenreduktion, Effizienzsteigerung)

Bedeutung der Value Proposition

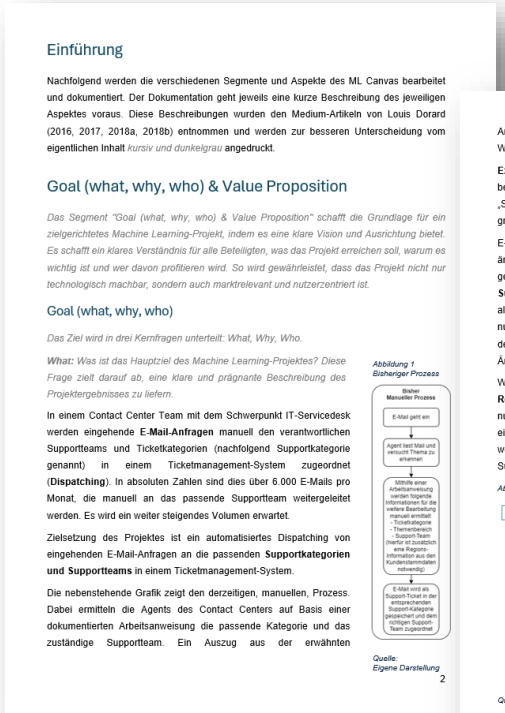
- Sicherheit in der Projektentscheidung
  „Warum machen wir das eigentlich?“
- Kommunikationswerkzeug für Stakeholder

Tipps zur Entwicklung einer starken Value Proposition

- Engen Austausch mit Stakeholdern / Projektauftraggebern pflegen
- Iteratives Feedback einholen und einarbeiten



3.1 Value Proposition - Beispiel



Einführung

Nachfolgend werden die verschiedenen Segmente und Aspekte des ML Canvas bearbeitet und dokumentiert. Der Dokumentation geht jeweils eine kurze Beschreibung des jeweiligen Aspektes voraus. Diese Beschreibungen wurden den Medium-Artikeln von Louis Dorard (2016, 2017, 2018a, 2018b) entnommen und werden zur besseren Unterscheidung vom eigentlichen Inhalt *kursiv und dunkelgrau* angedruckt.

Goal (what, why, who) & Value Proposition

Das Segment "Goal (what, why, who) & Value Proposition" schafft die Grundlage für ein zielgerichtetes Machine Learning-Projekt, indem es eine klare Vision und Ausrichtung bietet. Es schafft ein klares Verständnis für alle Beteiligten, was das Projekt erreichen soll, warum es wichtig ist und wer davon profitieren wird. So wird gewährleistet, dass das Projekt nicht nur technologisch machbar, sondern auch marktrelevant und nutzerzentriert ist.

Goal (what, why, who)

Das Ziel wird in drei Kernfragen unterteilt: What, Why, Who.

***What:** Was ist das Hauptziel des Machine Learning-Projektes? Diese Frage zielt darauf ab, eine klare und prägnante Beschreibung des Projektergebnisses zu liefern.*

In einem Contact Center Team mit dem Schwerpunkt IT-Servicedesk werden eingehende **E-Mail-Anfragen** manuell den verantwortlichen Supportteams und Ticketkategorien (nachfolgend Supportkategorie genannt) in einem Ticketmanagement-System zugeordnet (**Dispatching**). In absoluten Zahlen sind dies über 6.000 E-Mails pro Monat, die manuell an das passende Supportteam weitergeleitet werden. Es wird ein weiter steigendes Volumen erwartet.

Zielsetzung des Projektes ist ein automatisiertes Dispatching von eingehenden E-Mail-Anfragen an die passenden **Supportkategorien und Supportteams** in einem Ticketmanagement-System.

Die nebenstehende Grafik zeigt den derzeitigen, manuellen, Prozess. Dabei ermitteln die Agents des Contact Centers auf Basis einer dokumentierten Arbeitsanweisung die passende Kategorie und das zuständige Supportteam. Ein Auszug aus der erwähnten

Abbildung 1 Bisheriger Prozess

Quelle: Eigene Darstellung

Arbeitsanweisung ist weiter unten im Abschnitt „Learn – Data Sources" als Screenshot eines Wiki-Systems zu sehen.

Exkurs – Klärung von Begriffen und Zusammenhängen: Der nachfolgende Abschnitt beschreibt den Gesamtkontext des Dispatching-Prozesses aus Sicht der Dimensionen „Supportkategorie, Themenbereich, Supportteam und Region" und schließt mit einer grafischen Übersichtsdarstellung ab.

E-Mails werden zunächst in einer vom Auftraggeber vorgegebenen und sich sehr selten ändernden Struktur von **Supportkategorien** eingeordnet und anschließend von den Agents gedanklich nochmals einer Art „Unter-**Themenbereich**" zugeordnet, um das zuständige **Supportteam** zu ermitteln. Eine Besonderheit hierbei ist, dass diese **Themenbereiche** nicht als explizit zuordenbare Unterkategorie im Ticketsystem existieren, sondern von den Agents nur gedanklich dort eingeordnet werden, um das passende Supportteam aus einer Tabelle mit den dokumentierten Zuständigkeiten zu entnehmen. In diesem Bereich gibt es häufig Änderungen der Themenbereiche und der damit verbundenen Verantwortlichkeiten.

Weiterhin ist für die Auswahl des passenden **Supportteams** notwendig, die verantwortliche **Region** aus den Stammdaten des jeweiligen Kunden zu ermitteln, da die Supportteams nicht nur Themenabhängig zugeordnet sind, sondern auch regionsabhängig. Dies ist mit einem einfachen Mapping möglich, muss aber als Teil der Vorhersage der ML-Lösung mit abgedeckt werden. Es wird erwartet, dass lediglich die passende Supportkategorie und das passende Supportteam an das Ticketsystem zurückgegeben werden.

Abbildung 2 - Schematische Darstellung von Supportkategorien, Themenbereichen und Supportteams

Quelle: Eigene Abbildung

***Why:** Warum wird dieses Projekt durchgeführt? Hier geht es um die Begründung des Projekts, die sowohl geschäftliche als auch technische, soziale oder wirtschaftliche Beweggründe*

umfassen kann. Das „Warum" hilft dabei, die Bedeutung und die Dringlichkeit des Projekts zu verstehen.

- Mitarbeiterengpässe durch Arbeitskräftemangel
- Kostendruck durch Near-Shore und Off-Shore-Contact Center
- Wettbewerbsdruck durch Anbieter von automatisierter E-Mail-Bearbeitung

***Who:** Wer wird von diesem Projekt profitieren? Diese Frage klärt, wer die Zielgruppe ist, sei es ein internes Team, die Endkunden eines Unternehmens oder eine spezifische gesellschaftliche Gruppe. Die Identifizierung der Stakeholder ist entscheidend für die Ausrichtung des Projektes und die spätere Akzeptanz der Lösung.*

- Bestehendes Operations Team im Contact Center (Reduzierung der Engpässe und Überlastung)
- Key Account Manager und Sales (Angebot eines wettbewerbsfähigen Preises)
- Auftraggeber (Moderne, innovative und kosteneffiziente Lösung)

Value Proposition

Das Wertversprechen fasst zusammen, welchen einzigartigen Wert das Projekt seinen Stakeholdern bietet.

Pro manueller Bearbeitung geht man von einer Dauer zwischen 0,5 und 1 Minuten aus. Monatlich werden mindestens 6.000 E-Mails bearbeitet.

Zielsetzung ist, dass 80 % aller Anfragen künftig vollautomatisch richtig kategorisiert und den passenden Supportteam zugeordnet werden. Kann nur die Kategorie eindeutig vorhergesagt werden aber nicht das Supportteam, bedeutet dies für die Mitarbeiter dennoch eine Zeitersparnis, da sie so schneller den richtigen Themenbereich und Supportteam eingrenzen können.

Das Potenzial einer jährlichen Zeitersparnis wird dabei wie folgt ermittelt:

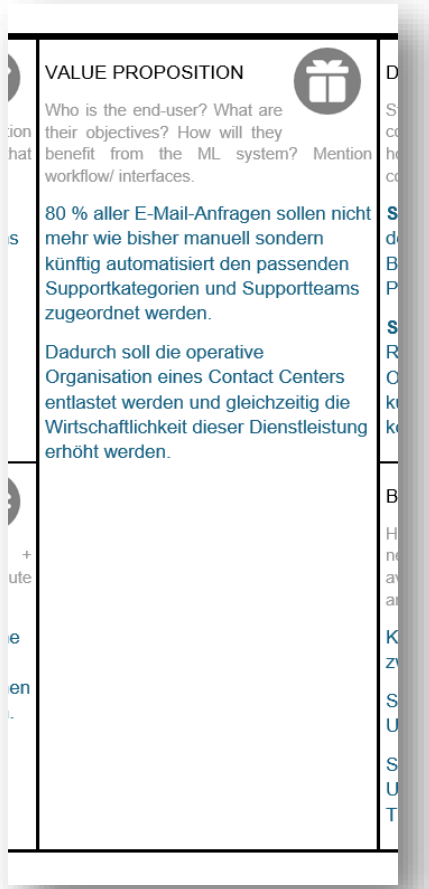
Best Guess: 6.000 Mails * 80 % * 1 Minute * 12 Monate = 57.600 Min. bzw. 960 Std./Jahr

Real Guess: 6.000 Mails * 80 % * 0.5 Minute * 12 Monate = 25.800 Min. bzw. 480 Std./Jahr

Bei einem kalkulatorischen Stundensatz von 45 € Vollkosten ergibt sich daraus ein Ersparnispotenzial von Mitarbeiterressourcen zwischen 21.600 € und 43.200 € pro Jahr.

Learn (how)

In diesem Segment des ML Canvas werden die Rahmenbedingungen ermittelt, wie eine mögliche Machine Learning Lösung die relevanten Daten und Informationen erhält, sodass ein passendes Modell trainiert werden kann.



VALUE PROPOSITION

Who is the end-user? What are their objectives? How will they benefit from the ML system? Mention workflow/ interfaces.

80 % aller E-Mail-Anfragen sollen nicht mehr wie bisher manuell sondern künftig automatisiert den passenden Supportkategorien und Supportteams zugeordnet werden.

Dadurch soll die operative Organisation eines Contact Centers entlastet werden und gleichzeitig die Wirtschaftlichkeit dieser Dienstleistung erhöht werden.

SAMHAMMER
The Service Excellence Company

3.2 Learn

Data Sources

- Welche Rohdatenquellen können genutzt werden?
- Wie zugänglich sind diese Daten?
- Was davon ist relevant für die Aufgabenstellung?

Collecting Data

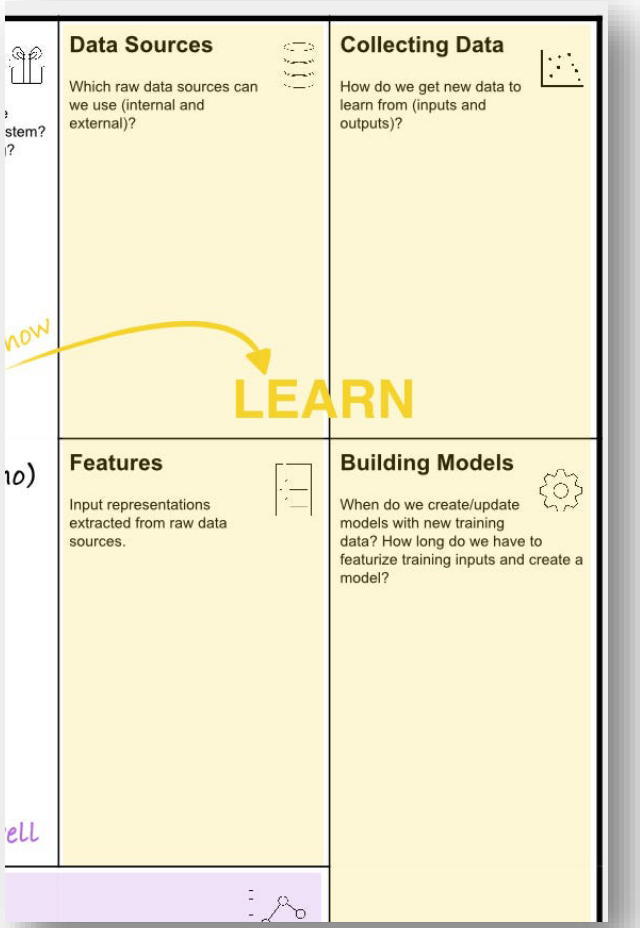
- Wie können die relevanten Daten gesammelt und gespeichert werden?
- Datenerfassungsintervalle?
- Art und Weise wie Daten gesammelt und ggf. angereichert werden (z.B. manuelles nachträgliches labeln oder Export bestehender Daten)

Features

- Feature-Auswahl: Identifikation der relevantesten und nützlichsten Features aus dem ursprünglichen Datensatz, die für das Modell am informativsten sind. Beispielsweise die E-Mail von Kunden, die Absenderadresse, die Betreffzeile und die Kundengruppe

Building Models

- Wie und wann soll das Modell mit aktuellen Trainingsdaten aktualisiert werden
- Aktualisierungszyklen von Modellen (Balance zwischen Ressourcenaufwand und Notwendigkeit, um gute Qualität zu erzielen)
- Abhängigkeit zu aktuellen und relevanten Trainingsdaten
- Wechselwirkung mit Collecting Data → Wie können alte oder falsche Daten entfernt werden?



3.2 Learn - Beispiel

Data Sources

Hier werden alle Ursprünge der für das Projekt ben
interne Datenbanken, öffentliche Datensätze, APIs
und IoT-Geräte generierte Daten. Auf dieser Stufe
von Daten verfügbar sind, wie zugänglich sie sind
Problem sind.

Supportkategorien

Historische Tickets aus Ticketmanagement
Diese enthalten die eingegangene E-Mail sow
Die Daten sind über eine SQL-Datenbank
Die Supportkategorien haben einen eher stabilen Ch

Supportteams (Definition des Themenbereichs

Das passende Supportteam ergibt sich anhand
Themenbereich (einer Art Unter-Einheit der zuvo
zuständigen Supportregion des Kunden.

Der Themenbereich wird den Tickets nicht explizi
Supportteam wäre als Datengrundlage vorhanden.
Themenbereiche verantwortlich sein. Es gibt hi
Supportteams, weshalb historische Daten nicht
Aufgrund kurzfristiger Änderungen in den Support
zudem möglich sein, geänderte Zuordnungen und
bei der Vorhersage zu berücksichtigen (quasi-Realti

Für die Beschreibung der Themenbereiche sow
existieren zwei Varianten (Variante 1a/1b und Varia

- **Datenquelle-Variante 1a – Unstrukturiert**
einem internen Wiki-System existiert
Beschreibung und Sammlung von Schlüssel
Nennung eines oder mehrerer passender Su
könnte beispielsweise als .pdf-Dokument z
einer ersten Sichtung wird dieses Dokument
Learning gesehen.

Abbildung 3 - Auszug aus Wiki-System Dokumentation von Themenbereichen und Supportteams

Quelle: Screenshot Wiki-System USU

Datenquelle-Variante 1b – Teilstrukturierte Dokum
Rahmen der Projektvorbereitungen wurden die Informat
einer objektorientierten Datenbank strukturiert und
Analyse- und Strukturierungs-Tool „Assistant Builde
Informationen sind analog der Variante 1a, jedoch ist die
vermuten, dass eine Anwendung für einen ML-Use Ca
Hinweis: Ein exemplarischer Auszug der Theme
Supportkategorie ist in **Anhang II** zu finden.

Abbildung 4 - Auszug aus teilstrukturierter Dokumentation von Them

Quelle: Screenshot "Assistant Builder"

- **Datenquelle-Variante 2 – Erzeugen von Trainingsdaten:** Falls keine der dargestellten Varianten 1a und 1b genutzt werden kann, um eine Vorhersage zu erzeugen, müssen zunächst Trainingsdat
einzelnen Themenbereiche benannt und
zugeordnet werden (Labeln von Daten).
analog den Ticketdaten aus dem Datawar

Data Collection

Dieser Aspekt befasst sich mit den Prozessen u
Speicherung der benötigten Daten eingesetzt
Datenpipelines, die Auswahl von Speich
Datenerfassungsintervallen einschließen. Die Art
hat direkte Auswirkungen auf die Qualität und Ver
Analyse und Modellierung.

Supportkategorien

Da sich die Struktur und Zuordnung zu den versch
hierfür historische Ticketdaten verwendet werden.

Das Datawarehouse kann optional angebunden we
ermöglichen. Da ein Training vermutlich eher selte
Jahr), können die Trainingsdaten auch manuell im

Kontaktgründe und Supportteams

Da die oben beschriebene Variante 1b (Objektdat
vermutlich zu bevorzugen ist, wird zunächst die Ve

Das führende System für die Dokumentation der
beschriebene Objektdatenbank (MongoDB). Id
Anbindung, die jederzeit nach einer Anpassun
Aktualisierte Fassung z.B. via REST API zur Verfü

Alternativ wäre denkbar, dass die ML Anwendung
Daten über eine REST API oder einen VPN-Tunn
der aktuellen Daten angestoßen werden kann (pull

Eine Live-Anbindung ist nicht notwendig, da nicht f
Objektdatenbank geladen werden müssen.

Sollte für die Vorhersage der Supportteams auf die
Themenbereichen) zurückgegriffen werden müsse
und Vorhersageprozess notwendig:

- Historische Ticketdaten (analog der Supportkategorien)
- Matrix, die eine Zuordnung von Themenbereichen und Supportteams je Region darstellt. Hierfür könnte ebenfalls die u
werden, die so angebunden werden
zuständigen Supportteams möglichst sc
Mapping muss zusätzlich im Ticketmana

Features

Dieser Schritt umfasst die Identifizierung, Auswa
Rohdaten, die für das zu lernende Modell am au
der Datenanalyse und -bereinigung sowie der Fe
Input-Daten in eine Form zu bringen, die von Alg

Vorhersage der Supportkategorie

Für die Vorhersage der Supportkategorie ist
ausreichend. Die Mitarbeiter im Contact Cente
ausschließlich die E-Mail des Kunden als Grund

Vorhersage des Supportteams

Für die Vorhersage des Supportteams wird nebe
den **Stammdaten des Kunden** genutzt, die zug

Im Falle eines zweistufigen Vorhersageprozesse
kann die im ersten Schritt ermittelte Supportkat
sein, da es auch Supportkategorien gibt, die nur

Building Models

"Building Models" umfasst Überlegungen darü
Trainingsdaten aktualisiert werden soll, um se
*Kernpunkte sind. **Aktualisierungszyklen** v*
Anwendungsbereich und der Dynamik des zu m
Balance zwischen Ressourcenaufwand und
*reagieren, gefunden werden. **Datenaktualität***
Trainingsdaten sind essenziell. Es kann vorteilh
Modellgenauigkeit und dessen Fähigkeit, die
verbessern.

Bei den **Supportkategorien** gibt es kaum Verär
sich von Zeit zu Zeit die von den Kunden gen
Applikationen und die Problembeschreibungen

Training alle 3 bis 6 Monate auf Basis der letzten 12 Monate empfohlen. Sollten sich dennoch Änderungen ergeben und Kategorien aufgelöst oder neue angelegt werden, ist auch dies der Auslöser für ein Training, bei dem die nicht mehr gültigen Datensätze ausgeschlossen werden müssen. Ein weiterer Auslöser für ein Training kann zudem eine Verschlechterung der Vorhersage-Ergebnisse sein, die über ein entsprechendes Monitoring erkannt wird. Ein Training wird sehr wahrscheinlich manuell ausgelöst und es erfolgt kein automatisch ausgelöstes Neu-Training.

Bei den Supportteams gibt es häufig Änderungen. Hier ist es notwendig, dass die Trainingsdaten möglichst aktuell sind bzw. nach einer Änderung, die es in der Regel monatlich gibt, sehr zeitnah ein Training stattfindet, sobald ausreichende Trainingsdaten verfügbar sind (Bei Training anhand historischer Daten) oder unmittelbar nach der Änderung neue Informationen für die Zuordnung der passenden Supportteams genutzt werden, sobald diese vorliegen (zum Beispiel bei One-Shot-Classification über LLMs durch Aktualisierung des Prompt-Templates).

Predict (how)

Der Abschnitt "Predict (how)" befasst sich damit, wie genau die Vorhersage (Prediction) im Kontext des Projekts gemacht werden soll. Ein zentraler Punkt beim Entwurf von ML-basierten Lösungen ist, zu verstehen, welche Art von Vorhersagen das System machen soll und wie diese Vorhersagen aus vorhandenen Daten abgeleitet werden können.

Decisions

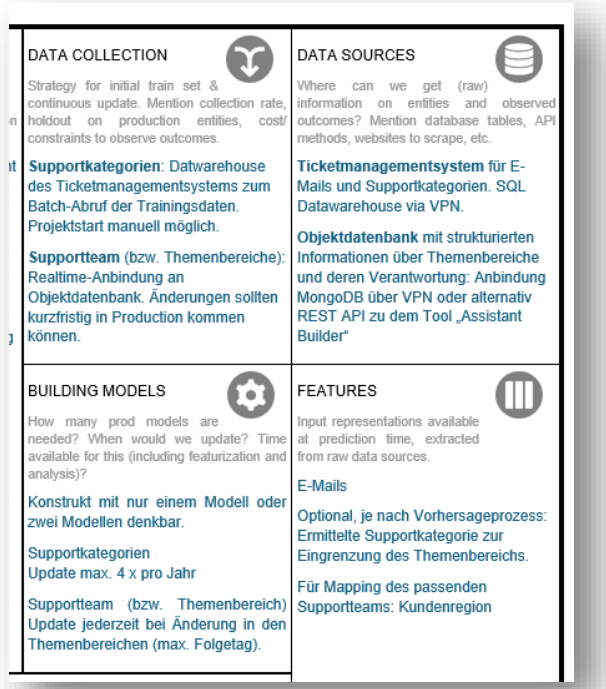
Der Aspekt „Decisions" liefert eine detaillierte Beschreibung dessen, was das System vorhersagen soll, also die Entscheidungen, die auf Grundlage der Vorhersagen des Modells getroffen werden. Es geht darum zu verstehen, wie die Ausgaben (Predictions) des Modells in praktische, wertsteigernde Entscheidungen innerhalb einer Anwendung oder eines Geschäftsprozesses umgesetzt werden können. Dies beinhaltet die Überlegung, welche Handlungen auf Basis der Vorhersagen ergriffen werden und wie diese Handlungen zur Erreichung der übergeordneten Ziele beitragen.

Für einen automatisierten Arbeitsablauf ist es notwendig, die Supportkategorie und das Supportteam automatisiert vorherzusagen.

Abhängig von der Modellauswahl und Modellqualität, sind zwei verschiedene Ansätze denkbar, die nachfolgend als Alternativen Decisions-V1 und Decisions-V2 beschrieben sind.

Decisions-V1 – zweistufiger Prozess: Vorhersage der Supportkategorie und nachgelagerte Vorhersage des Themenbereichs mit entsprechendem Mapping zum passenden Supportteam

9



DATA COLLECTION

Strategy for initial train set & continuous update. Mention collection rate, holdout on production entities, cost/ constraints to observe outcomes.

Supportkategorien: Datwarehouse des Ticketmanagementsystems zum Batch-Abruf der Trainingsdaten. Projektstart manuell möglich.

Supportteam (bzw. Themenbereiche): Realtime-Anbindung an Objektdatenbank. Änderungen sollten kurzfristig in Production kommen können.

DATA SOURCES

Where can we get (raw) information on entities and observed outcomes? Mention database tables, API methods, websites to scrape, etc.

Ticketmanagementsystem für E-Mails und Supportkategorien. SQL Datawarehouse via VPN.

Objektdatenbank mit strukturierten Informationen über Themenbereiche und deren Verantwortung: Anbindung MongoDB über VPN oder alternativ REST API zu dem Tool „Assistant Builder"

BUILDING MODELS

How many prod models are needed? When would we update? Time available for this (including featurization and analysis)?

Konstrukt mit nur einem Modell oder zwei Modellen denkbar.

Supportkategorien
Update max. 4 x pro Jahr

Supportteam (bzw. Themenbereich) Update jederzeit bei Änderung in den Themenbereichen (max. Folgetag).

FEATURES

Input representations available at prediction time, extracted from raw data sources.

E-Mails

Optional, je nach Vorhersageprozess: Ermittelte Supportkategorie zur Eingrenzung des Themenbereichs.

Für Mapping des passenden Supportteams: Kundenregion

SAMHAMMER
The Service Excellence Company

3.2 Learn - Beispiel

Data Sources are Sources of Knowledge!
Keep their Quality in Mind!

SAMHAMMER
The Service Excellence Company

3.3 Predict

Decisions

- Detaillierte Beschreibung: Was genau soll das Modell vorhersagen?
- Wann soll die Vorhersage stattfinden?
- Verständnis erlangen: Welche Handlungen werden auf Basis des Modells ausgelöst?

ML Task / Prediction Task

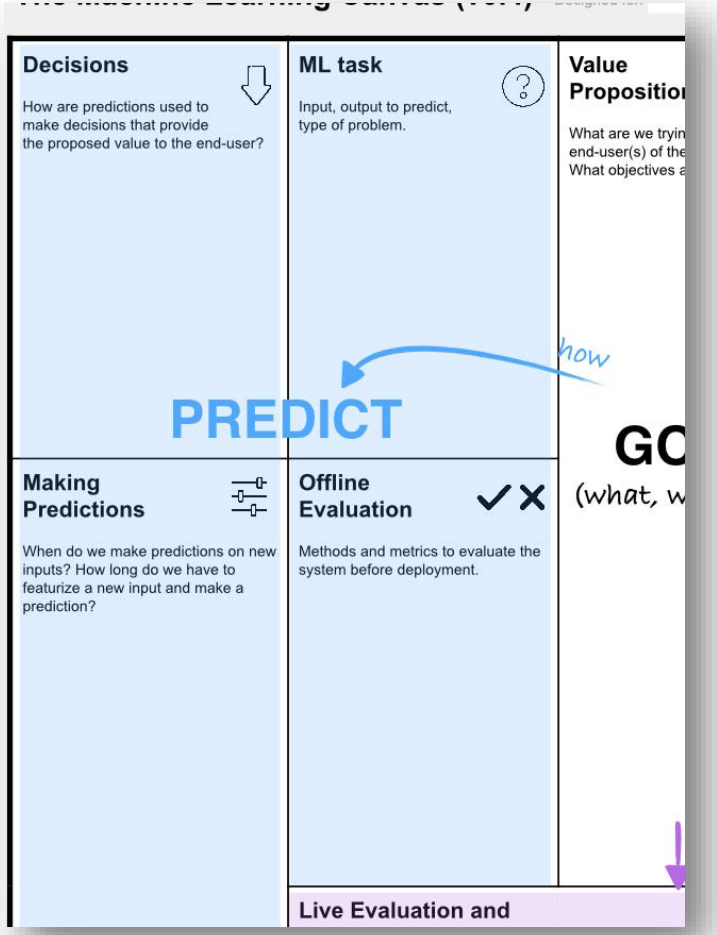
- Welche ML Aufgabe soll gelöst werden?
  (z.B Klassifikation, Regression, Clustering, Anomalieerkennung ….)

Making Predictions

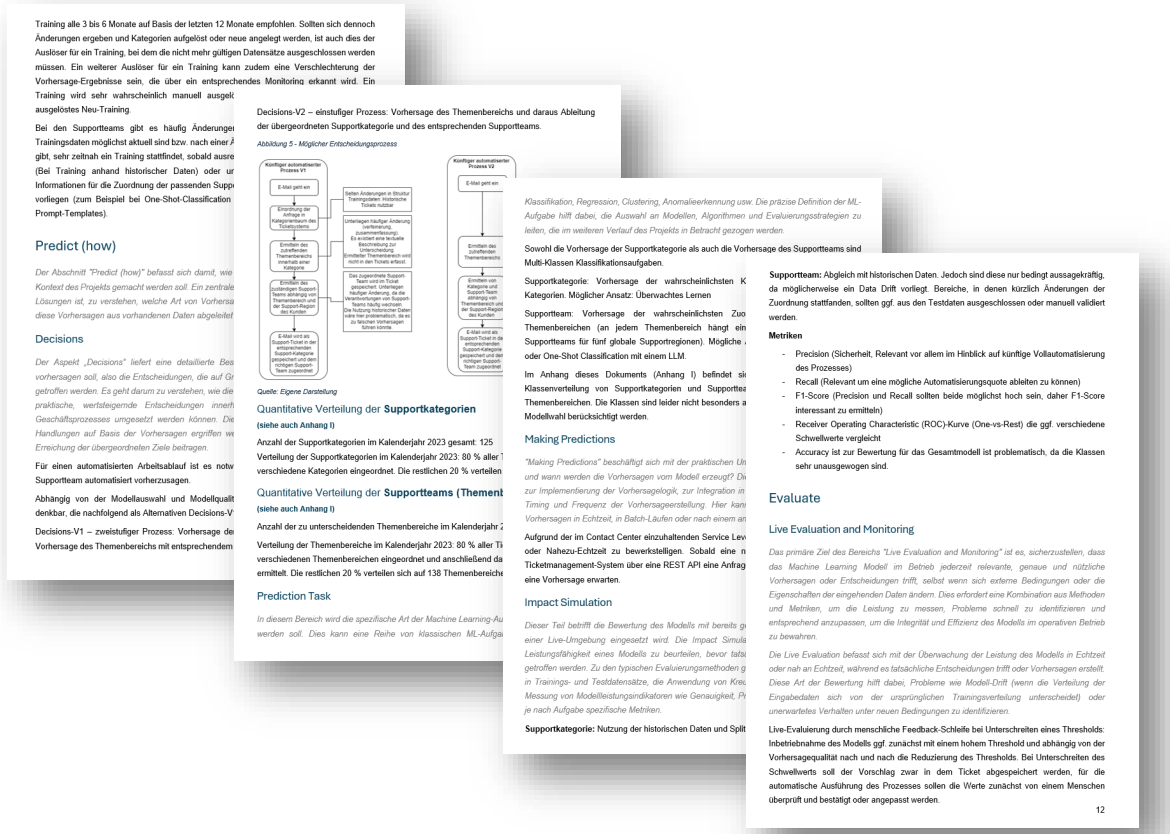
- Wie und wann werden die Vorhersagen vom Modell erzeugt?
- Timing und Frequenz der Vorhersage (z.B. RealTime vs. Batch)
- Notwendige Schnittstellen zur Anbindung / Nutzung von REST APIs / …
- Reaktionszeiten und Datenströme

Offline Evaluation / Impact Simulation

- Bewertung des Modells bevor es in Produktion geht
- Mindestqualität des Modells
- Fokus auf Maximierung der Vorhersagemenge vs. Vorhersagequalität (Metriken wie Recall, Precision, F1-Score,…)
- Testdatensätze



3.3 Predict - Beispiel



Training alle 3 bis 6 Monate auf Basis der letzten 12 Monate empfohlen. Sollten sich dennoch Änderungen ergeben und Kategorien aufgelöst oder neue angelegt werden, ist auch dies der Auslöser für ein Training, bei dem die nicht mehr gültigen Datensätze ausgeschlossen werden müssen. Ein weiterer Auslöser für ein Training kann zudem eine Verschlechterung der Vorhersage-Ergebnisse sein, die über ein entsprechendes Monitoring erkannt wird. Ein Training wird sehr wahrscheinlich manuell ausgelöst ausgelöstes Neu-Training.

Bei den Supportteams gibt es häufig Änderungen Trainingsdaten möglichst aktuell sind bzw. nach einer gibt, sehr zeitnah ein Training stattfindet, sobald ausre (Bei Training anhand historischer Daten) oder ur Informationen für die Zuordnung der passenden Supp vorliegen (zum Beispiel bei One-Shot-Classification Prompt-Templates).

Predict (how)

Der Abschnitt "Predict (how)" befasst sich damit, wie Kontext des Projekts gemacht werden soll. Ein zentrale Lösungen ist, zu verstehen, welche Art von Vorhersa diese Vorhersagen aus vorhandenen Daten abgeleitet

Decisions

Der Aspekt „Decisions" liefert eine detaillierte Bes vorhersagen soll, also die Entscheidungen, die auf Gr getroffen werden. Es geht darum zu verstehen, wie die praktische, wertsteigernde Entscheidungen innerh Geschäftsprozesses umgesetzt werden können. Die Handlungen auf Basis der Vorhersagen ergriffen w Erreichung der übergeordneten Ziele beitragen.

Für einen automatisierten Arbeitsablauf ist es notw Supportteam automatisiert vorherzusagen.

Abhängig von der Modellauswahl und Modellqualit denkbar, die nachfolgend als Alternativen Decisions-V

Decisions-V1 – zweistufiger Prozess: Vorhersage de Vorhersage des Themenbereichs mit entsprechendem

Decisions-V2 – einstufiger Prozess: Vorhersage des Themenbereichs und daraus Ableitung der übergeordneten Supportkategorie und des entsprechenden Supportteams.

Abbildung 5 - Möglicher Entscheidungsprozess

Quelle: Eigene Darstellung

Quantitative Verteilung der **Supportkategorien** (siehe auch Anhang I)

Anzahl der Supportkategorien im Kalenderjahr 2023 gesamt: 125

Verteilung der Supportkategorien im Kalenderjahr 2023: 80 % aller T verschiedene Kategorien eingeordnet. Die restlichen 20 % verteilen

Quantitative Verteilung der **Supportteams (Themen** (siehe auch Anhang I)

Anzahl der zu unterscheidenden Themenbereiche im Kalenderjahr 2

Verteilung der Themenbereiche im Kalenderjahr 2023: 80 % aller Ti verschiedenen Themenbereichen eingeordnet und anschließend da ermittelt. Die restlichen 20 % verteilen sich auf 138 Themenbereiche

Prediction Task

In diesem Bereich wird die spezifische Art der Machine Learning-Au werden soll. Dies kann eine Reihe von klassischen ML-Aufga

Klassifikation, Regression, Clustering, Anomalieerkennung usw. Die präzise Definition der ML-Aufgabe hilft dabei, die Auswahl an Modellen, Algorithmen und Evaluierungsstrategien zu leiten, die im weiteren Verlauf des Projekts in Betracht gezogen werden.

Sowohl die Vorhersage der Supportkategorie als auch die Vorhersage des Supportteams sind Multi-Klassen Klassifikationsaufgaben.

Supportkategorie: Vorhersage der wahrscheinlichsten K Kategorien. Möglicher Ansatz: Überwachtes Lernen

Supportteam: Vorhersage der wahrscheinlichsten Zuo Themenbereichen (an jedem Themenbereich hängt ein Supportteams für fünf globale Supportregionen). Mögliche oder One-Shot Classification mit einem LLM.

Im Anhang dieses Dokuments (Anhang I) befindet si Klassenverteilung von Supportkategorien und Supportte Themenbereichen. Die Klassen sind leider nicht besonders a Modellwahl berücksichtigt werden.

Making Predictions

"Making Predictions" beschäftigt sich mit der praktischen Üb und wann werden die Vorhersagen vom Modell erzeugt? Die zur Implementierung der Vorhersagelogik, zur Integration in Timing und Frequenz der Vorhersageerstellung. Hier kann Vorhersagen in Echtzeit, in Batch-Läufen oder nach einem an

Aufgrund der im Contact Center einzuhaltenden Service Lev oder Nahezu-Echtzeit zu bewerkstelligen. Sobald eine n Ticketmanagement-System über eine REST API eine Anfrag eine Vorhersage erwarten.

Impact Simulation

Dieser Teil befrifft die Bewertung des Modells mit bereits g einer Live-Umgebung eingesetzt wird. Die Impact Simula Leistungsfähigkeit eines Modells zu beurteilen, bevor tatsä getroffen werden. Zu den typischen Evaluierungsmethoden g in Trainings- und Testdatensätze, die Anwendung von Kre Messung von Modellleistungsindikatoren wie Genauigkeit, Pr je nach Aufgabe spezifische Metriken.

Supportkategorie: Nutzung der historischen Daten und Split

Supportteam: Abgleich mit historischen Daten. Jedoch sind diese nur bedingt aussagekräftig, da möglicherweise ein Data Drift vorliegt. Bereiche, in denen kürzlich Änderungen der Zuordnung stattfanden, sollten ggf. aus den Testdaten ausgeschlossen oder manuell validiert werden.

Metriken

- Precision (Sicherheit, Relevant vor allem im Hinblick auf künftige Vollautomatisierung des Prozesses)
- Recall (Relevant um eine mögliche Automatisierungsquote ableiten zu können)
- F1-Score (Precision und Recall sollten beide möglichst hoch sein, daher F1-Score interessant zu ermitteln)
- Receiver Operating Characteristic (ROC)-Kurve (One-vs-Rest) die ggf. verschiedene Schwellwerte vergleicht
- Accuracy ist zur Bewertung für das Gesamtmodell ist problematisch, da die Klassen sehr unausgewogen sind.

Evaluate

Live Evaluation and Monitoring

Das primäre Ziel des Bereichs "Live Evaluation and Monitoring" ist es, sicherzustellen, dass das Machine Learning Modell im Betrieb jederzeit relevante, genaue und nützliche Vorhersagen oder Entscheidungen trifft, selbst wenn sich externe Bedingungen oder die Eigenschaften der eingehenden Daten ändern. Dies erfordert eine Kombination aus Methoden und Metriken, um die Leistung zu messen, Probleme schnell zu identifizieren und entsprechend anzupassen, um die Integrität und Effizienz des Modells im operativen Betrieb zu bewahren.

Die Live Evaluation befasst sich mit der Überwachung der Leistung des Modells in Echtzeit oder nah an Echtzeit, während es tatsächliche Entscheidungen trifft oder Vorhersagen erstellt. Diese Art der Bewertung hilft dabei, Probleme wie Modell-Drift (wenn die Verteilung der Eingabedaten sich von der ursprünglichen Trainingsverteilung unterscheidet) oder unerwartetes Verhalten unter neuen Bedingungen zu identifizieren.

Live-Evaluierung durch menschliche Feedback-Schleife bei Unterschreiten eines Thresholds: Inbetriebnahme des Modells ggf. zunächst mit einem hohem Threshold und abhängig von der Vorhersagequalität nach und nach die Reduzierung des Thresholds. Bei Unterschreiten des Schwellwerts soll der Vorschlag zwar in dem Ticket abgespeichert werden, für die automatische Ausführung des Prozesses sollen die Werte zunächst von einem Menschen überprüft und bestätigt oder angepasst werden.

12

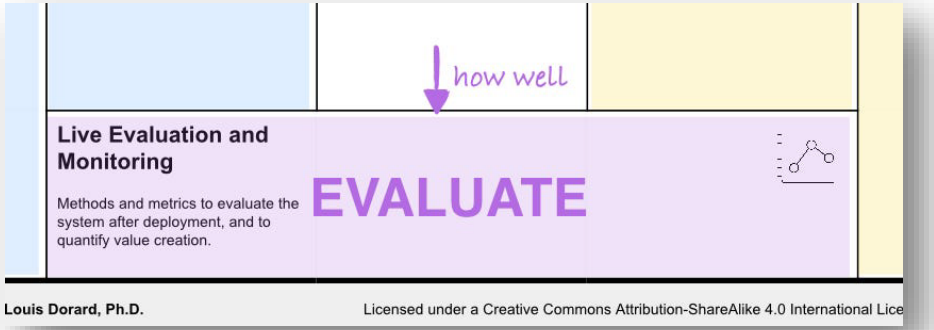
PREDICTION TASK	DECISIONS
Type of task? Entity on which predictions are made? Possible outcomes? Wait time before observation? Supportkategorien: Realtime Multi-Klassen Klassifikation Supportteams (bzw. Themenbereiche: Realtime Multi-Klassen Klassifikation Möglicherweise auch One-Shot Classification via LLM statt klassischem überwachten Training bei den Themenbereichen.	How are predictions turned into proposed value for the end-user? Mention parameters of the process / application that does that. Automatische Vorhersage von Supportkategorien und Supportteams (bzw. den Themenbereichen) zum automatischen Dispatching von E-Mails an das zuständige Bearbeitungsteam.
IMPACT SIMULATION	**MAKING PREDICTIONS**
Can models be deployed? Which test data to assess performance? Cost/gain values for (in)correct decisions? Fairness constraint? Precision Recall F1-Score Receiver Operating Characteristic (ROC)-Kurve (One-vs-Rest) die verschiedene Schwellwerte vergleicht	When do we make real-time / batch pred.? Time available for this + featurization + post-processing? Compute target? Vorhersagen sollten nahezu Realtime erfolgen um die E-Mails schnellstmöglich bearbeiten zu können und die vereinbarten SLAs zu halten.

SAMHAMMER
The Service Excellence Company

3.4 Evaluate

Live Evaluation and Monitoring

- **Zielsetzung:** Sicherstellen, dass das Machine Learning Modell im Betrieb jederzeit relevante, genaue und nützliche Vorhersagen oder Entscheidungen trifft
- Definition von Methoden und Metriken, um
  - Leistung zu messen,
  - Probleme schnell zu identifizieren,
  - Integrität und Effizienz des Modells im operativen Betrieb zu bewahren.
- Überwachung der Leistung des Modells in Echtzeit oder nah an Echtzeit
- Zielgruppe und notwendige Automatismen



3.3 Evaluate - Beispiel

Supportteam: Abgleich mit historischen Daten. Jedoch sind diese nur bedingt aussagekräftig, da möglicherweise ein Data Drift vorliegt. Bereiche, in denen kürzlich Änderungen der Zuordnung stattfanden, sollten ggf. aus den Testdaten ausgeschlossen oder manuell validiert werden.

Metriken

- Precision (Sicherheit, Relevant vor allem im Hinblick auf künftige Vollautomatisierung des Prozesses)
- Recall (Relevant um eine mögliche Automatisierungsquote ableiten zu können)
- F1-Score (Precision und Recall sollten beide möglichst hoch sein, daher F1-Score interessant zu ermitteln)
- Receiver Operating Characteristic (ROC)-Kurve (One-vs-Rest) die ggf. verschiedene Schwellwerte vergleicht
- Accuracy ist zur Bewertung für das Gesamtmodell ist problematisch, da die Klassen sehr unausgewogen sind.

Evaluate

Live Evaluation and Monitoring

Das primäre Ziel des Bereichs "Live Evaluation and Monitoring" ist es, sicherzustellen, dass das Machine Learning Modell im Betrieb jederzeit relevante, genaue und nützliche Vorhersagen oder Entscheidungen trifft, selbst wenn sich externe Bedingungen oder die Eigenschaften der eingehenden Daten ändern. Dies erfordert eine Kombination aus Methoden und Metriken, um die Leistung zu messen, Probleme schnell zu identifizieren und entsprechend anzupassen, um die Integrität und Effizienz des Modells im operativen Betrieb zu bewahren.

Die Live Evaluation befasst sich mit der Überwachung der Leistung des Modells in Echtzeit oder nah an Echtzeit, während es tatsächliche Entscheidungen trifft oder Vorhersagen erstellt. Diese Art der Bewertung hilft dabei, Probleme wie Modell-Drift (wenn die Verteilung der Eingabedaten sich von der ursprünglichen Trainingsverteilung unterscheidet) oder unerwartetes Verhalten unter neuen Bedingungen zu identifizieren.

Live-Evaluierung durch menschliche Feedback-Schleife bei Unterschreiten eines Thresholds: Inbetriebnahme des Modells ggf. zunächst mit einem hohem Threshold und abhängig von der Vorhersagequalität nach und nach die Reduzierung des Thresholds. Bei Unterschreiten des Schwellwerts soll der Vorschlag zwar in dem Ticket abgespeichert werden, für die automatische Ausführung des Prozesses sollen die Werte zunächst von einem Menschen überprüft und bestätigt oder angepasst werden.

12

Kontinuierliches Monitoring der Metriken analog zur Impact Simulation (Precision, Recall, F1-Score). Im Ticketmanagement-System werden die Predictions sowie die endgültig gewählten Supportkategorien und Supportteams abgespeichert. Auf dieser Basis können die entsprechenden Metriken ermittelt und Handlungsfelder schnell ermittelt werden.

13

EVALUATION / MONITORING

Metrics to quantify value creation and measure the ML system's impact in production (on end-users and business)?

Precision, Recall, F1-Score

Aus Business-Sicht: Voll automatisierte Vorgänge, Teilautomatisierte Vorgänge mit Menschlichem Feedback und falsche Vorhersagen im Verhältnis zum Gesamtvolumen.

SAMHAMMER
The Service Excellence Company

4. Hinweise zur Nutzung

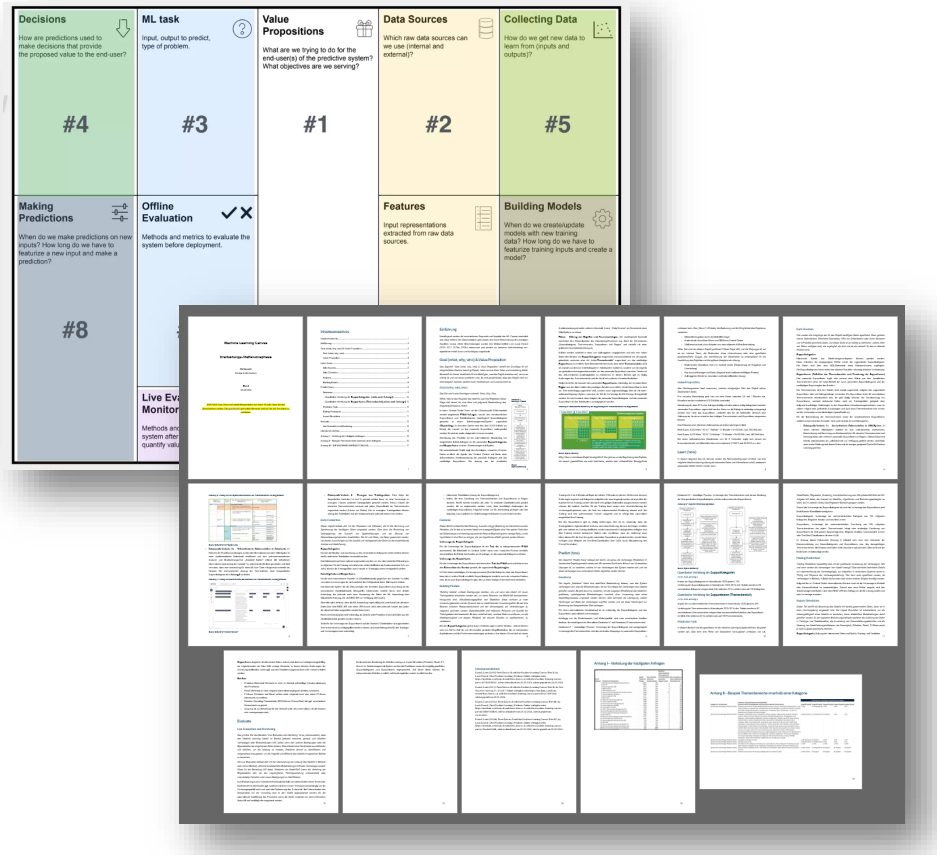
Die einzelnen Elemente können als eine Art Checkpunkt verstanden werden
Empfehlung zur Reihenfolge!

Das Canvas ist als eine Aggregation zu verstehen.

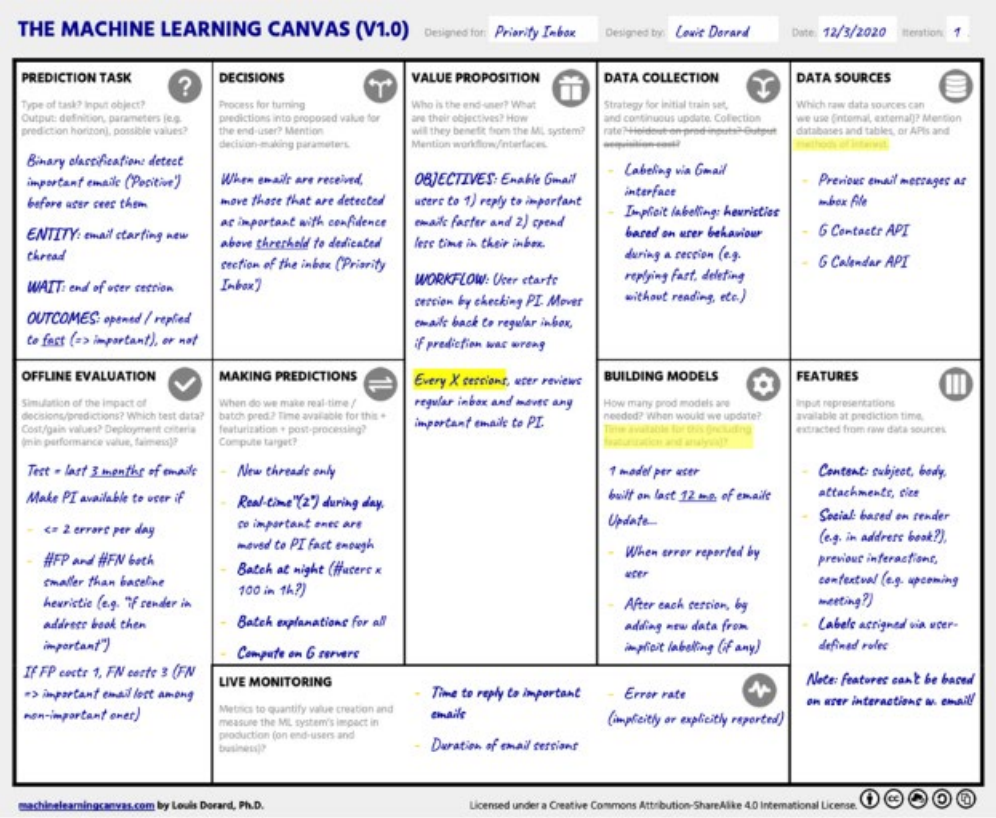
Der zentrale Wert liegt in der resultierenden ausführlichen Dokumentation.

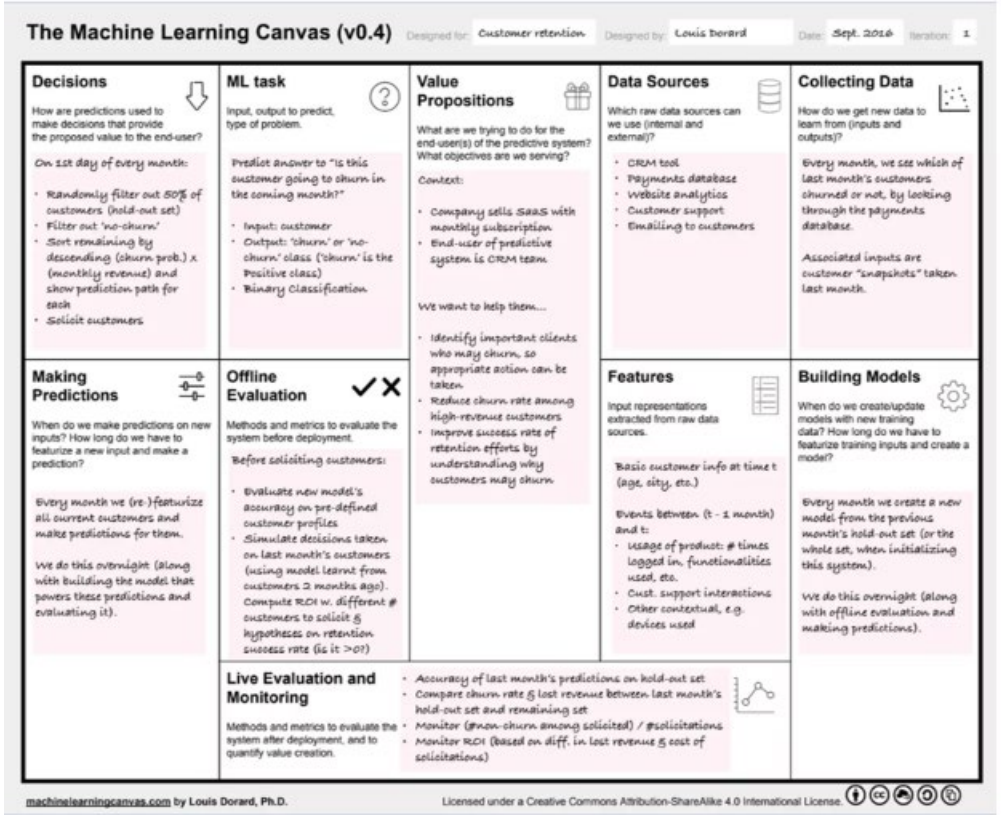
Gap: Bewertung der inhaltlichen Datenqualität vor Projektstart wird nicht explizit behandelt!

Tipp: Die Anwender sollten durch einen kontinuierlich wachsenden Katalog mit Hinweisen für die einzelnen Module unterstützt werden.



4. Hinweise zur Nutzung – Weitere Beispiele





Quelle: Louis Dorard, machinelearningcanvas.com

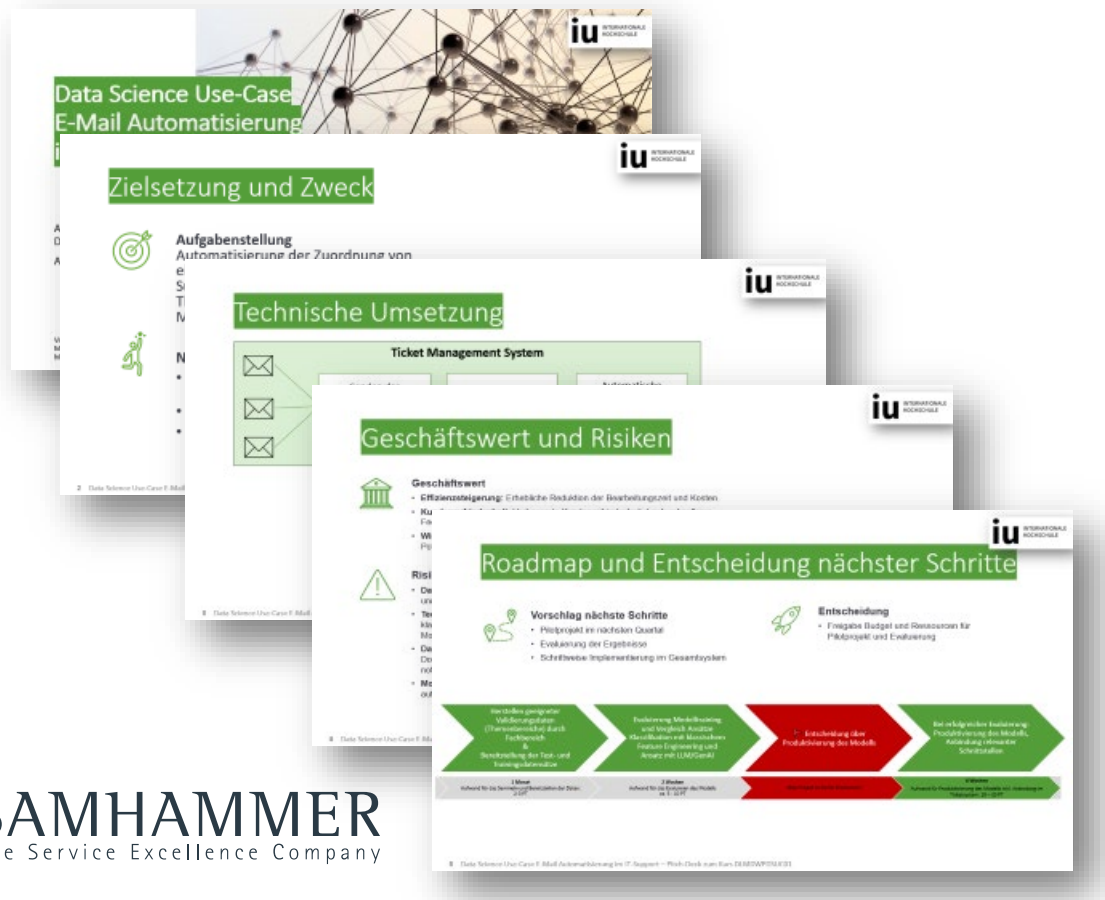
5. Persönliches Fazit zum ML Canvas

Durch kontinuierlich wachsenden UseCase-Katalog entsteht ein wertvolles Asset für Ihr Unternehmen

Wiederholbare und strukturierte Vorgehensweise
→ Gemeinsamer Anker in Projekten

Ergebnisse können verdichtet für Entscheidungsvorlagen und Management Summary wiederverwendet werden

Reviewprozess der Dokumentation durch Projektunbeteiligte ist bei der Einführung sehr hilfreich



SAMHAMMER
The Service Excellence Company

Entscheidungsvorlage basierend auf Erkenntnissen und Dokumentation des angewendeten ML Canvas

Und Jetzt?

Zeit für gemeinsame Diskussion und Austausch

Monika Arbter-Hubrich,
IU Internationale Hochschule & Samhammer AG