

„Anwendungsfelder der Business Analytics am Beispiel von CAD-Systemen innerhalb der Radiologie“

Marc-David Krain // 527373

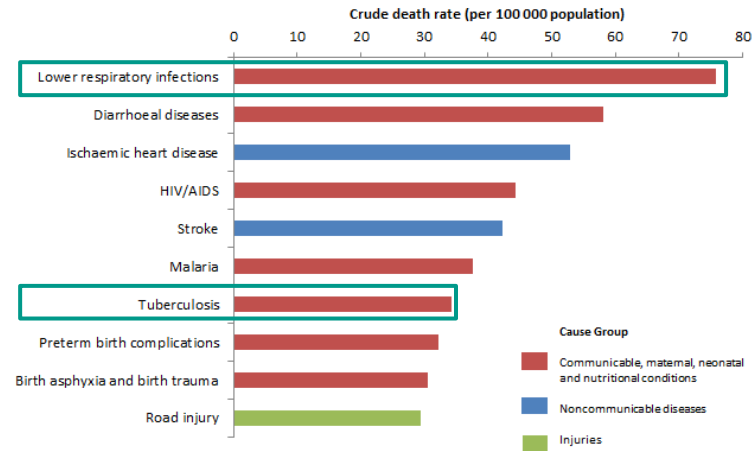
Inhalt

1. Ausgangssituation
 2. Anwendungsfall & Funktionsbereich
 3. Problem- & Zielstellung des Projektes
 4. Forschungsfragen des Projektes
 5. In Scope vs. Out-of-Scope
 6. Vorgehensweise gemäß Zielstellung
 7. Potenzielle Restriktionen
 8. Vorläufige Gliederung der Ausarbeitung
 9. Verwendete Literatur
-

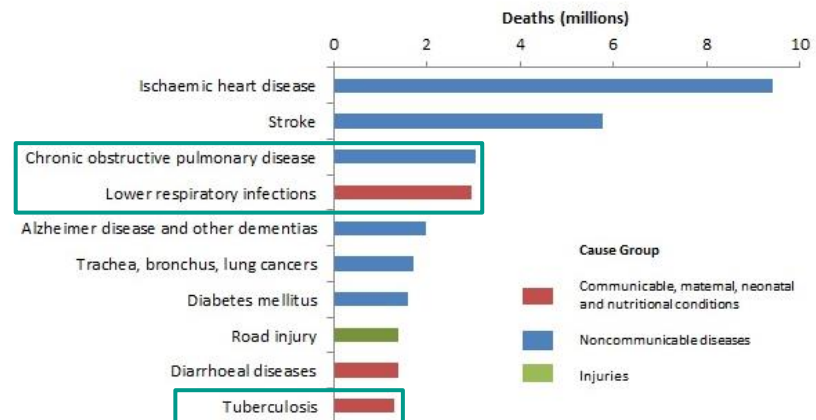
Ausgangssituation – Respiratorische Erkrankungen

Respiratorische Erkrankungen sind zusammengefasst eine der führenden Todesursachen innerhalb der modernen Gesellschaft

Top 10 causes of deaths in low-income countries in 2016



Top 10 global causes of deaths, 2016



Ausgangssituation - Pneumonie

Pneumonie als Breitbandindikator für Respiratorische Erkrankungen

Auslöser einer Lungenentzündung

Infektiöse Lungenentzündung

„Typische Lungenentzündung“



Bakterien

- Meist durch Pneumokokken als Tröpfcheninfektion übertragen

„Atypische Lungenentzündung“



Mykoplasmen

- Sehr kleine Bakterien, die häufig für Kleinerepidemien an Schulen o.Ä. verantwortlich sind



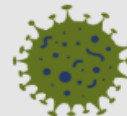
Legionellen

- Über verschmutztes (Trink-)Wasser übertragen



Chlamydien

- Übertragung als Tröpfcheninfektion
- Lösen auch eine Geschlechtskrankheit oder Augenentzündungen aus



Viren

- Entstehen häufig durch das Influenzavirus, das als Tröpfcheninfektion übertragen wird

Nicht-infektiöse Lungenentzündung



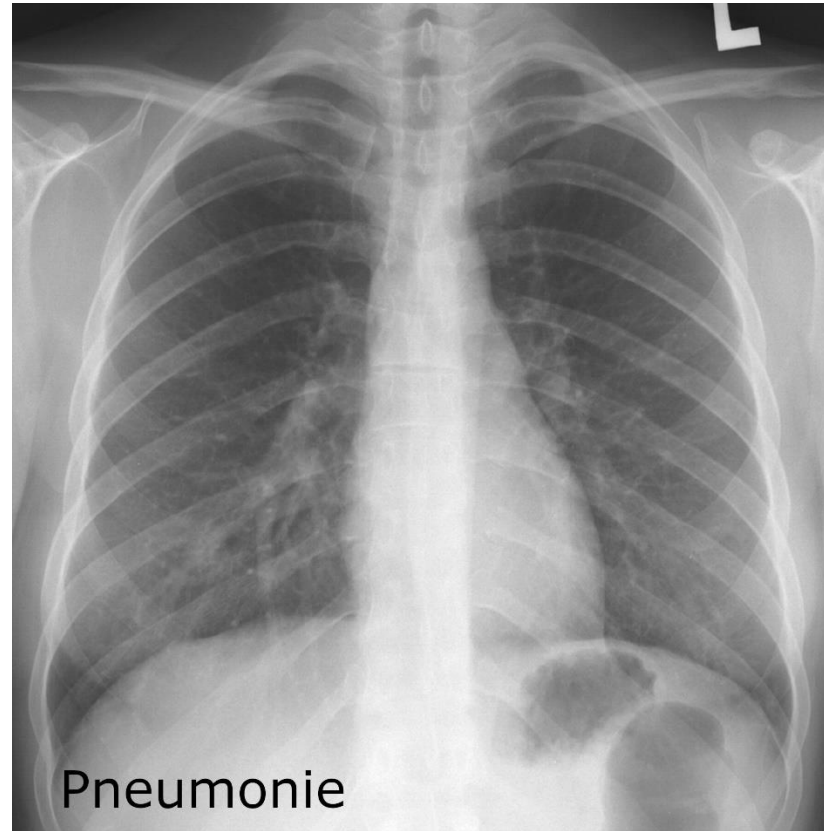
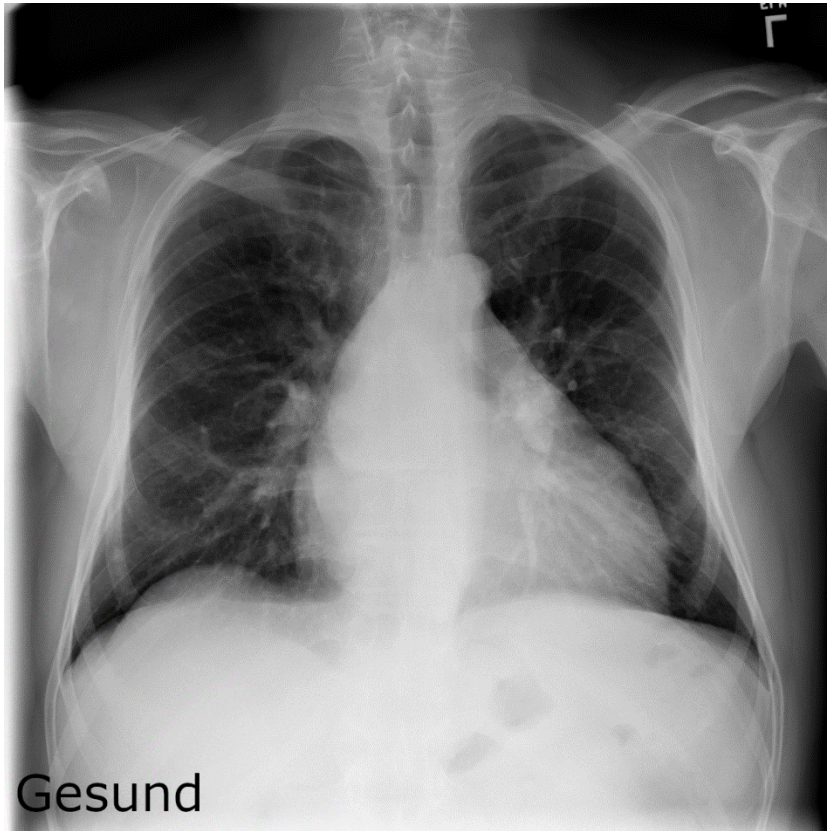
- Ausgelöst durch das Einatmen chemischer Gase, Staub oder Dämpfe

Aspirationspneumonie

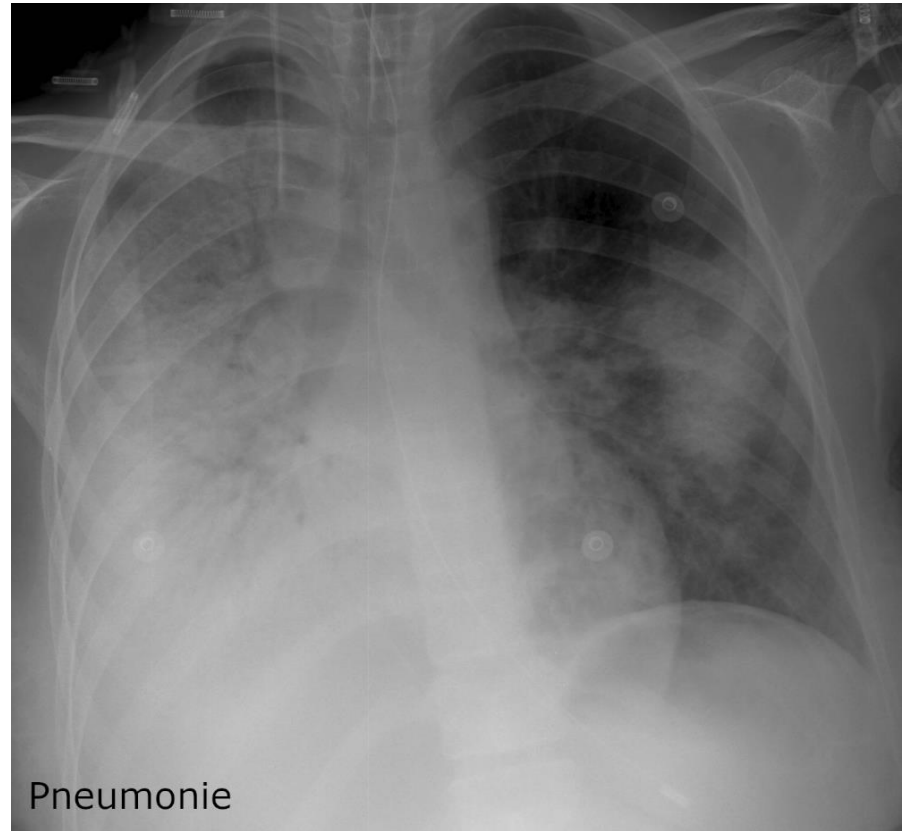
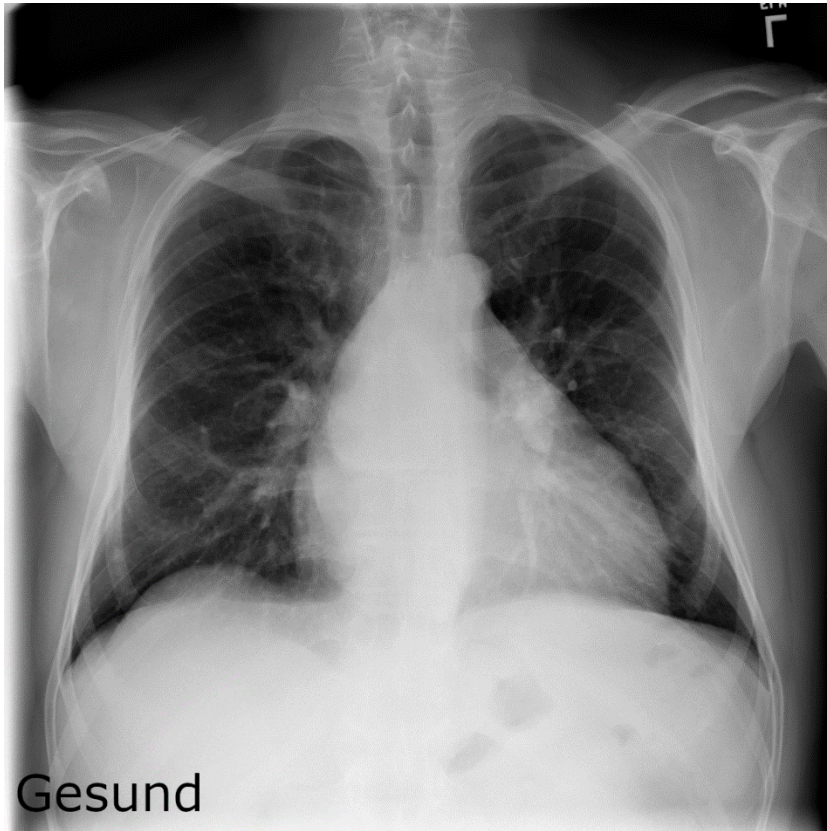


- Entsteht, wenn Nahrung, Flüssigkeit oder Magensäure eingeatmet („aspiriert“) wird und dadurch über die Luft-röhre in die Lunge gelangt

Ausgangssituation – Radiologie & Pneumonie

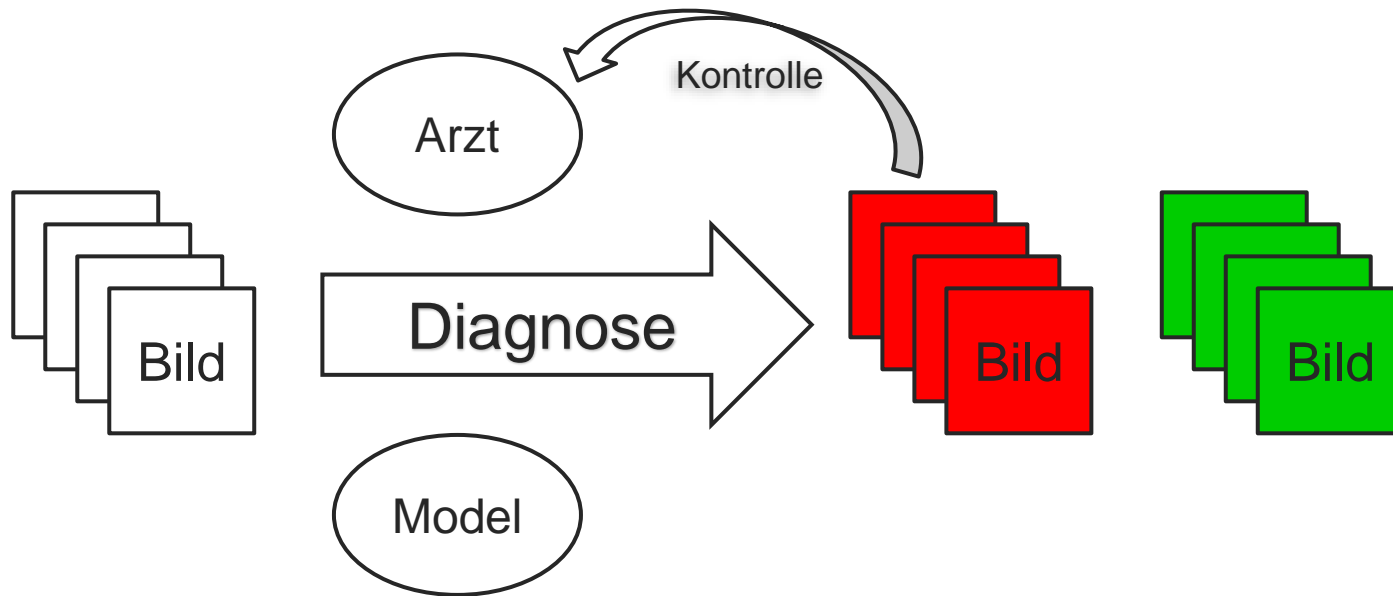


Ausgangssituation – Radiologie & Pneumonie



Ausgangssituation - Computer Aided Detection

“Computer aided detection (CAD) is a technology designed to decrease observational oversights and thus the false negative rates of physicians interpreting medical images.” [1][2]



Sensitivität in Radiologie[3]: 94% ➡ 98%

Sensitivität in Mammografie[4]: 74% ➡ 87%

Anwendungsfall

CAD-System für Röntgenaufnahmen zur Anomalieerkennung

Funktionsbereich

Pharmazeutisch/Medizinischer Hintergrund:

Data Science for Radiology & Medical Affairs

Problemstellung und Zielstellung

Problemstellung

- Pneumonie ist initialer Breitband-Indikator für potentiell schwerwiegende Krankheiten
- Entwicklungsländer haben erhöhtes Risikopotential
- Identifikation von Anomalien benötigt geschultes Personal und Zeit

Zielstellung

Ziel dieses Projektes ist die Entwicklung einer CAD-Anwendung, zur Unterstützung von medizinischen Personal, die die folgenden Kriterien erfüllt:

1. Open Source
2. Low/No Budget
3. Minimal Complexity

Forschungsfragen des Projektes

Zentrale Forschungsfrage des Projektes:

„Ist es möglich basierend auf öffentlich verfügbaren Datenquellen und Technologien ein verlässliches CAD-System zur Unterstützung von medizinischem Personal bereitzustellen?“

Untergeordnete Forschungsfragen des Projektes:

1. Sind öffentliche Röntgenbilder ausreichend für ein CAD?
2. Welche Algorithmen sind zielführend?
3. Welche Technologien und Ansätze sind zielführend?

In Scope vs. Out-of-Scope

In Scope

1. Konzipierung/Entwicklung von Modellen basierend auf Transfer Learning/Deep Learning
2. Entwicklung eines Minimal Viable Products anhand von ausgewähltem DS-Prozessmodell
3. Bereitstellung eines theoretischen Rollout & Industrialisierungskonzepts für spätere Arbeiten

Out of Scope

1. praktischer Rollout & Industrialisierung des Minimal Viable Products
 2. Bereitstellung von ausführlicheren Produktdokumentationen für Drittnutzer
-

Vorgehensweise – Datensätze für das Projekt

Quelle	URL	Anzahl / Größe
Kaggle	https://www.kaggle.com/nih-chest-xrays/data	112,120 Bilder / ~ 42 GB
Kaggle	https://www.kaggle.com/khongweihao/covid19-xray-dataset-train-test-sets	188 Bilder / ~ 85.24 MB
Kaggle	https://www.kaggle.com/bachhrr/covid-chest-xray?select=images	359 Bilder / ~ 243.52 MB
Kaggle	https://www.kaggle.com/asraf047/covid19-pneumonia-normal-chest-xray-pa-dataset	4575 Bilder / ~ 1.38 GB
Kaggle	https://www.kaggle.com/pcbreviglieri/pneumonia-xray-images	5856 Bilder / ~ 1.15 GB

Vorgehensweise – CRISP-DM

CRISP-DM

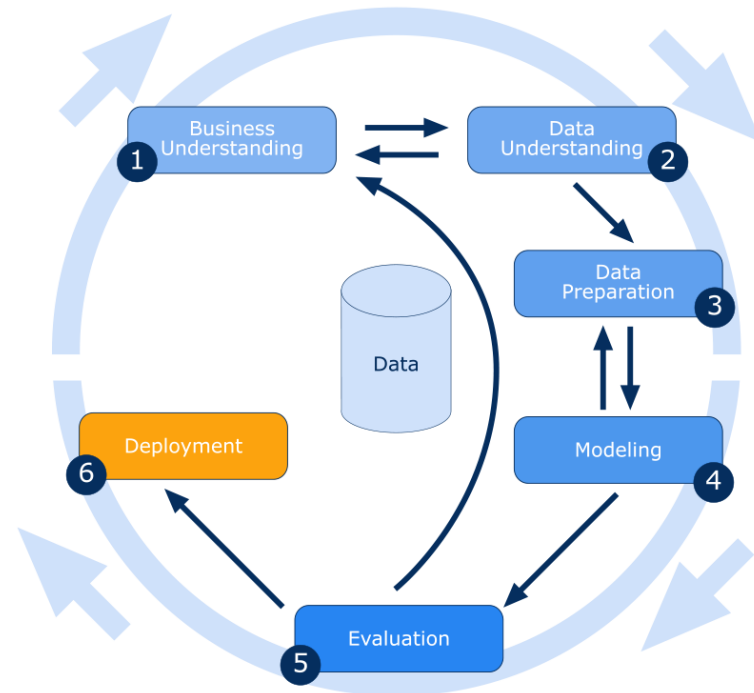
Pro:

Etabliertes Modell für DS-Projekte

Contra:

Fehlende Elemente im Vergleich zu ASUM-DM:

1. Zusammenarbeit
2. Versionskontrolle
3. Sicherheit
4. Compliance



Vorgehensweise – ASUM-DM

ASUS-DM

Pro:

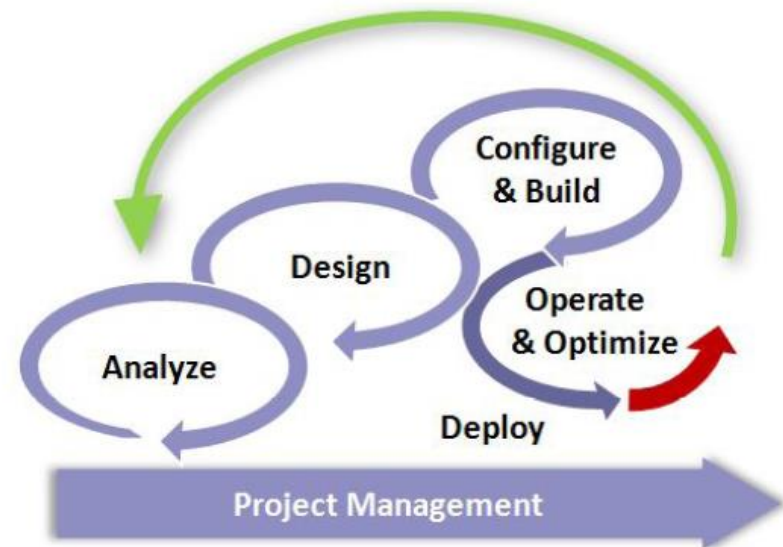
Prozess parallel zu CRISP-DM &

Neue Elemente:

1. Zusammenarbeit
2. Versionskontrolle
3. Sicherheit
4. Compliance

Contra:

Urtümlicher Fokus auf Agilität und Projektmanagement von Mehrpersonen-Team







































Vorgehensweise – Auswahl eines der genannten Werkzeuge

Kriterien für den Entscheidungsprozess:

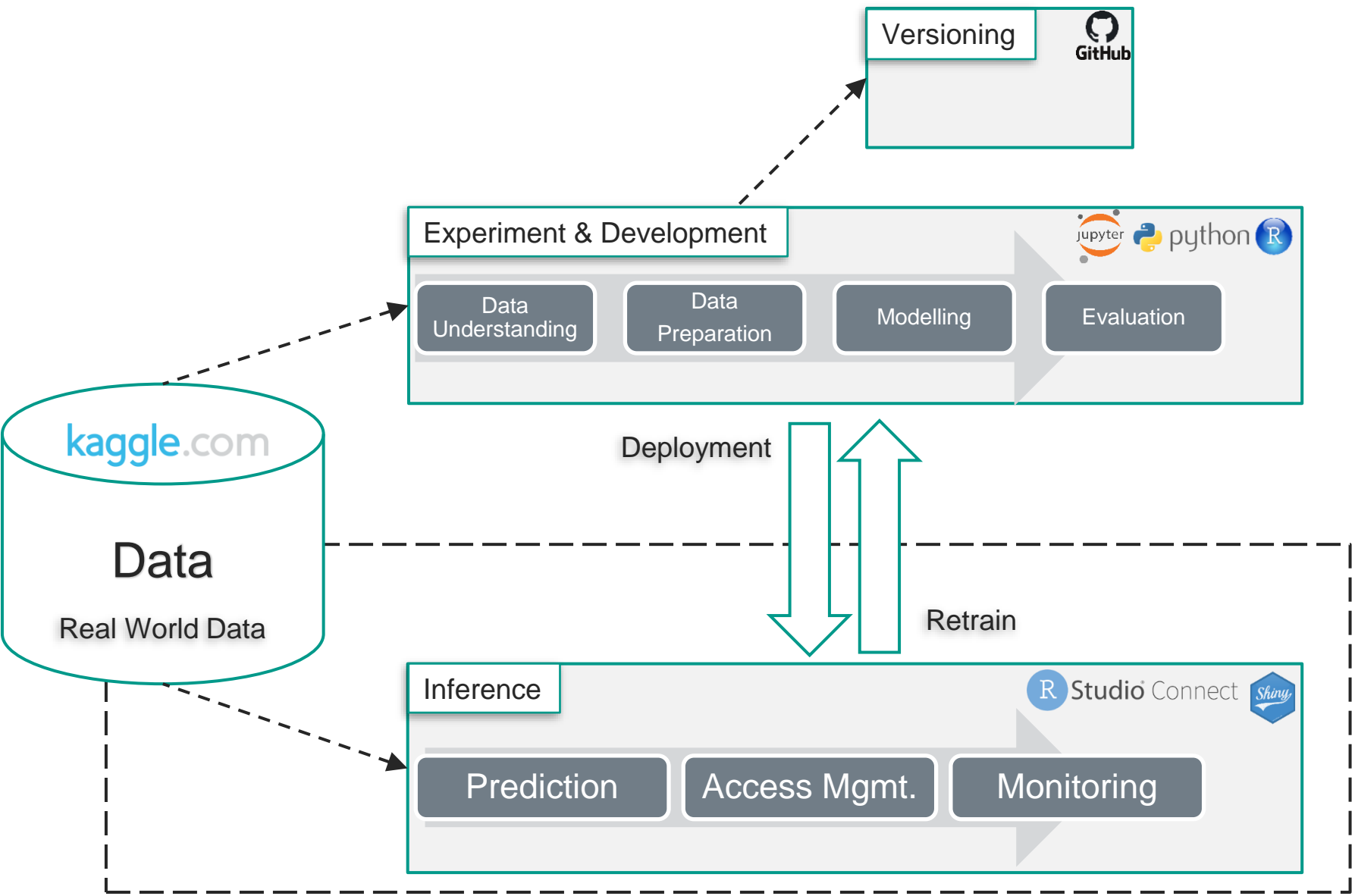
1. Bisherige Erfahrung
2. Open Source
3. Flexibilität
4. Vollständigkeit



Vorgehensweise – Nutzwertanalyse

Kriterium	%	1	2	3	4	5
Bisherige Erfahrung	40	   	 			
Open Source	30	     				  
Flexibilität	15	  	 	  		
Vollständigkeit	15	 	  	  		

Vorgangsweise - Workflow innerhalb des Projektes



Experimentelle Fragestellung bedarf erheblichen Mehraufwand zur Validierung:

- Ist GxP-Konformität ein relevanter Faktor?
- Datenschutzfragestellungen?
- Wie kann man Tests zur Erweisung der Wirksamkeit gestalten?
 - Implementierung und Partnerschaften in Praxis

Vorläufige Gliederung des Ausarbeitung

Titelblatt

Sperrvermerk

Verzeichnisse

Inhalt

Abkürzungen

Abbildungen

Tabellen

Symbole

Textteil

Einleitung

Methodik und
Werkzeuge

Implementierung

Ergebnis

Fazit

Zusammenfassung

Literaturverzeichnis

Anhang und
Appendum

Ehrenwörtliche
Erklärung

Verwendete Literatur - Onlinequellen

- [1] Castellino RA, Computer aided detection (CAD): an overview. *Cancer Imaging*, Published 2005 Aug 23, doi:10.1102/1470-7330.2005.0018
 - [2] Bhagirathi Halall et al, Computer Aided Diagnosis - Medical Image Analysis Techniques, Published: 2017 Dec 20, doi: 10.5772/intechopen.69792
 - [3] Zhimin Huo et al, Breast Cancer: Effectiveness of Computer-aided Diagnosis Observer Study with Independent Database of Mammograms, Published 2002 Aug 1, doi: 10.1148/radiol.2242010703
 - [4] Quek et al, Radiologists' detection of mammographic abnormalities with and without a computer-aided detection system, Published 2003 Aug 11, doi:10.1046/j.1440-1673.2003.01173.x
 - [5] <https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death>
 - [6] <https://www.intelligent-analysieren.de/news/crisp-dm-wird-zu-asum-dm.html>
-

Verwendete Literatur - Daten

<https://www.kaggle.com/nih-chest-xrays/data>

<https://www.kaggle.com/khoongweihao/covid19-xray-dataset-train-test-sets>

<https://www.kaggle.com/bachrr/covid-chest-xray?select=images>

<https://www.kaggle.com/asraf047/covid19-pneumonia-normal-chest-xray-pa-dataset>

<https://www.kaggle.com/pcbreviglieri/pneumonia-xray-images>

Verwendete Literatur - Bildquellen

Röntgenbilder: <https://www.kaggle.com/nih-chest-xrays/data>

Todesursachen: <https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death>

Pneumonie-Ursachen: <https://www.gesundheitswissen.de/innere-medicin/lungenentzuendung/>

CRISP-DM: <https://www.ifad.de/services/data-mining/anwendungsgebiete-des-data-mining/>

ASUM-DM: <https://statistik-dresden.de/archives/15783>

Verwendete Literatur - Bildquellen

R-Studio Logo: <https://www.pngegg.com/en/png-teymy/>

R-Shiny Logo: https://www.pngkey.com/detail/u2q8i1u2t4t4o0w7_r-shiny-logo/

KNIME-Logo: <https://www.abax.at/knime-datasciencelearnathon/>

Jupyter Logo:

https://upload.wikimedia.org/wikipedia/commons/thumb/3/38/Jupyter_logo.svg/518px-Jupyter_logo.svg.png

RS-Connect Logo:

<https://d33wubrfki0l68.cloudfront.net/83c6d3a12023cc739c8648629922557265fe34cf/44946/assets/img/connect-logo.png>

H2O.ai Logo [https://s3.amazonaws.com/owler-](https://s3.amazonaws.com/owler-image/logo/h2o_owler_20190611_121259_original.png)

[image/logo/h2o_owler_20190611_121259_original.png](https://s3.amazonaws.com/owler-image/logo/h2o_owler_20190611_121259_original.png)

Verwendete Literatur - Bildquellen

Tableau Logo: <https://www.absentdata.com/tableau-alternatives/attachment/tableau-logo/>

Python Logo: <https://www.pngegg.com/en/png-hcbwh>

QLIK Logo: <https://www.pngegg.com/en/png-zwxnh>

Rapidminer Logo: <https://academy.rapidminer.com/learning-paths/get-started-with-rapidminer-and-machine-learning>

Kaggle Logo: <https://logodix.com/kaggle>

SAS Logo: <https://logodix.com/logos/69503>

Alteryx Logo: <https://logodix.com/logos/449418>

Github Logo: https://www.vhv.rs/viewpic/hxbRimR_github-logo-png-github-transparent-png/
