INSTITUT FÜR DIGITALE GESUNDHEITSDATEN

Zusammenführen und Validieren von Krebsregisterdaten durch KI-Verfahren (ZuVaKI)

GMDS Jahrestagung

Dr. Nils Herm-Stapelberg // 09.09.2025 // Jena

Krebsregister Rheinland-Pfalz
im Institut für digitale Gesundheitsdaten

ZuVaKI - Ziele / Projektbeschreibung

*„Ermittlung der Potentiale von KI-Verfahren
zur Qualitätssicherung und
zum Erstellen von Best-Of-Datensätzen
in Krebsregistern“*

Untersuchungsfelder

- Identifizierung von auffälligen Datensätzen in Krebsregisterdaten (Anomalieerkennung)
- Zusammenführung von ggf. widersprüchlichen Informationen zu Tumorerkrankungen zu einer Best-Of-Tumorhistorie (Record Fusion)

Weitere Informationen:

- https://zuvaki.de
- Laufzeit: 2022 – 2026

ZuVaKI Partner

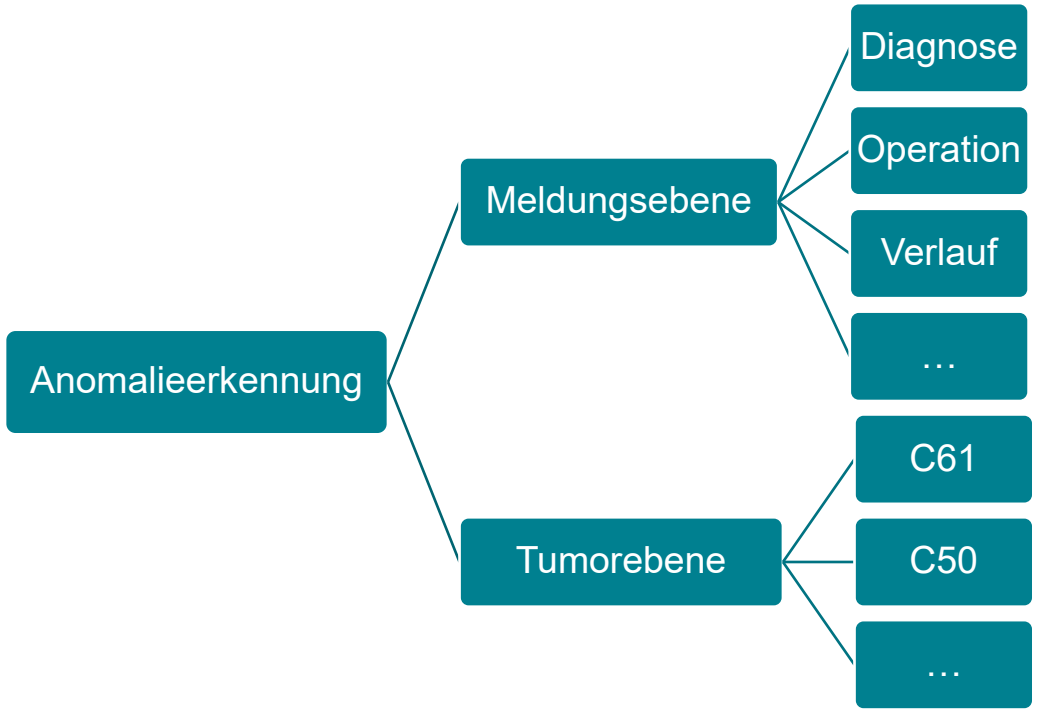
- Krebsregister Rheinland-Pfalz
- Klinisches Krebsregister Niedersachen
- Krebsregister Baden-Württemberg
- Hessisches Krebsregister
- Deutsches Kinderkrebsregister
- Johannes Gutenberg-Universität Mainz

Anomalieerkennung – Methode zur Qualitätssicherung

	Gezielte Qualitätssicherung Abfragen / Plausisbilitätsprüfungen	Ungezielte Qualitätssicherung Stichproben
Vorteile	• Finde nur tatsächliche Qualitätsprobleme • Art des Qualitätsproblems bekannt	• Finde Qualitätsprobleme ohne Einschränkungen • Finde Qualitätsprobleme mit komplexen Zusammenhängen
Limitationen	• Finde nur Qualitätsprobleme, nach denen ich suche • Komplexe Zusammenhänge sind schwer abzubilden	• Nicht alle Fälle sind Qualitätsprobleme • Art des Qualitätsproblems unbekannt

Anomalieerkennung

Anwendungsebenen



Vorgehen:

- Methoden: Unsupervised Learning
  - **FindFPOF (Frequent Pattern Outlier Factor)**
  - **Autoencoder**
- Output: keine Labels, sondern Anomaliescore, der den Grad der Anomalie widerspiegelt
- Evaluation: Fälle mit höchsten Anomaliescore wurden von Experten überprüft

Anomalieerkennung

Frequent Pattern Outlier Factor (FindFPOF)

Idee:

- Normale Meldungen haben viele häufige Muster
- Anomalien haben wenige häufige Muster

- Ein Muster ist eine Teilmenge von Werten, so dass es für jede Variable höchstens einen Wert gibt
- FPOF eines Datensatzes ist die Summe der Häufigkeiten aller übereinstimmenden häufigen Muster geteilt durch die Anzahl aller häufigen Muster

- R-Paket: “fpmoutliers”
- Gute Interpretierbarkeit!

Anomalieerkennung

(Denoising) Autoencoder

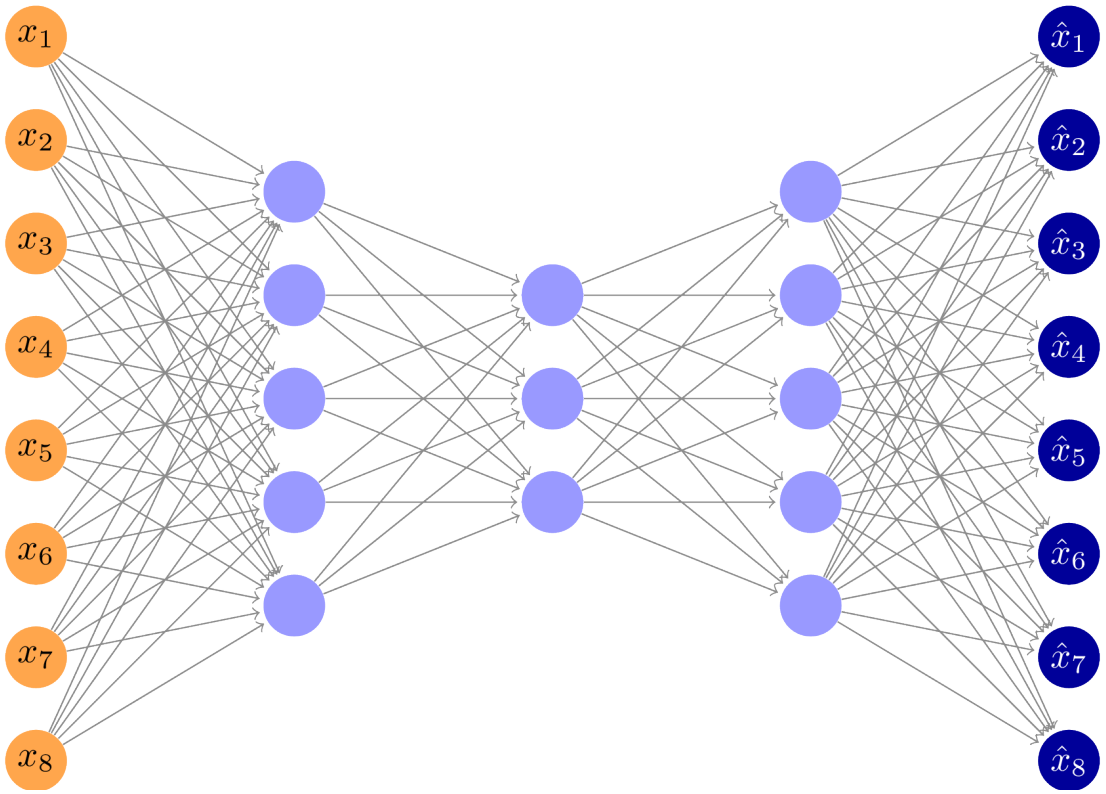
Input:

ICD-10: 50.9

ICD-O morphology: 8140/3

Grading: 2

…



Output:

ICD-10: 50.9

ICD-O morphology: 8140/3

Grading: T

…

Anomalieerkennung Meldungsebene

Ablauf Anomalieerkennung Meldungsebene:

1. Mit den Methoden FindFPOF und Autoencoder wird für jede Meldung je ein Anomaliescore erstellt
2. Die Meldungen mit den höchsten Scores werden zur Evaluation in einer Liste gespeichert (Top 100)
3. Medizinische Dokumentare/Dokumentarinnen schauen sich die Listen an und überprüfen, ob tatsächlich Anomalien vorliegen

Beispiele für erkannte Anomalien:

Melde-anlass	Meldung ID	Geschlecht	Alter Diagnose	Unter-suchungs-datum	Gesamt-beurteilung	Beurteilung Primärtumor	Beurteilung Lymphknoten	Beurteilung Fernmetas-tasen
Verlauf	xxxxx	M	73	04.06.2020	**P (Progression)**	**R (Rezidiv)**	X	X

Angabe Primärtumor und Gesamtbeurteilung passen nicht zusammen

Melde-anlass	Meldung ID	Geschlecht	Alter Diagnose	Beginn Datum	Intention	Gesamtdosis	Einzeldosis	Zielgebiet
Strahlen-therapie	xxxxx	W	58	11.03.2018	K	/	/	3.1

Therapieangaben fehlen

Anomalieerkennung Meldungsebene

Ergebnisse Evaluation Meldungsebene:

nach Meldeanlass

	Gesamt	Diagnose	Operation	Strahlen-therapie	System-therapie	Verlauf
KR1	**39,68 % (198)**	35,00 % (35)	9,00 % (9)	**77,78 % (77)**	15,15 % (15)	**61,39 % (62)**
KR2	**49,30 % (246)**	63,00 % (63)	21,78 % (22)	**75,00 % (75)**	28,28 % (28)	**58,59 % (58)**

nach Anomaliemethode

Krebs-register	**Anzahl Meldungen in Anomalieliste insgesamt**	**Schnittmenge Anomalieliste FPOF + Autoencoder**	**Anteil Anomalien insgesamt**	Anteil Anomalien unter den Meldungen mit höchstem **FPOF-Score**	Anteil Anomalien unter den Meldungen mit höchstem **Autoencoder-Score**
KR1	499 (280 je Methode)	61	39,68 % (198)	**40,00 % (112)**	**42,86 % (120)**
KR2	499 (258 je Methode)	17	49,30 % (246)	**52,33 % (135)**	**45,74 % (118)**

→ Unterschiedliche Methoden (FindFPOF, Autoencoder) finden unterschiedliche Anomalien, aber performen ähnlich

Anomalieerkennung Tumorebene

Vorgehen:

- Auf Tumorebene fand die Anomalieerkennung analog statt (FindFPOF + Autoencoder)
- Je Tumor wurden die Meldungen zu Variablen zusammengefasst, die den Tumor ganzheitlich beschreiben:
  - Anzahl Therapiemeldungen
  - Erste Therapie
  - Abstand Diagnose und Therapie in Tagen
  - …

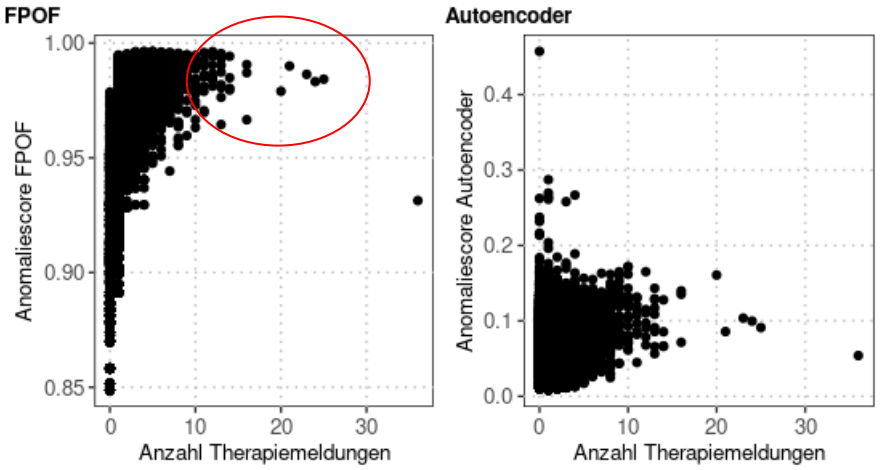
- **Betrachtete Entitäten:**
  - C61 Prostata
  - C50 Mamma
  - C33-34 Lunge
  - C91-95 Leukämie

Anzahl Tumoren:

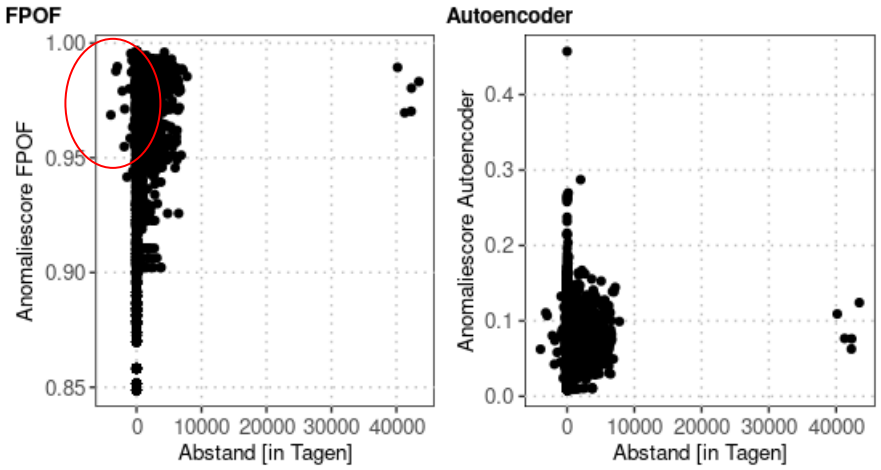
Projektpartner	C61	C50	C33 – C34	C91 – C95
Krebsregister 1	22.533	28.942	18.570	3.169
Krebsregister 2	25.955	27.923	16.041	2.917
Krebsregister 3	4.849	7.485	4.876	558

Anomalieerkennung – Tumorebene – C61

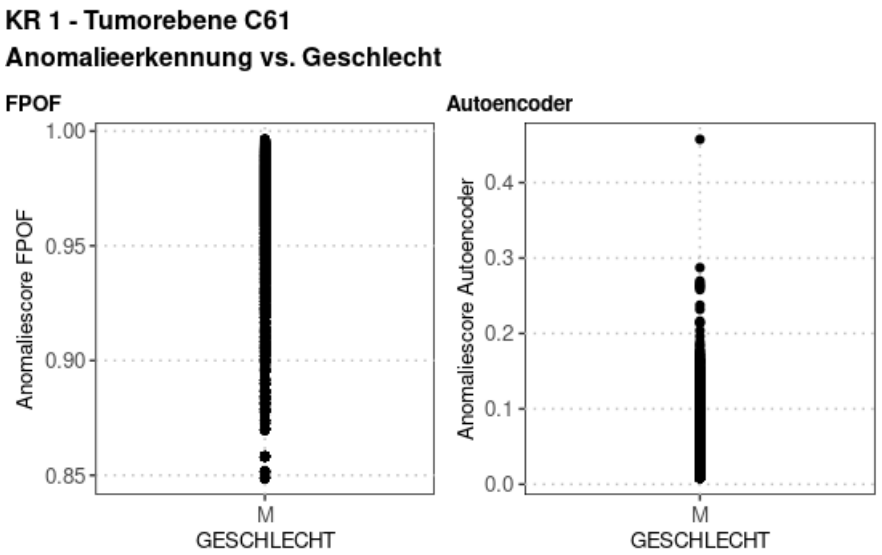
1



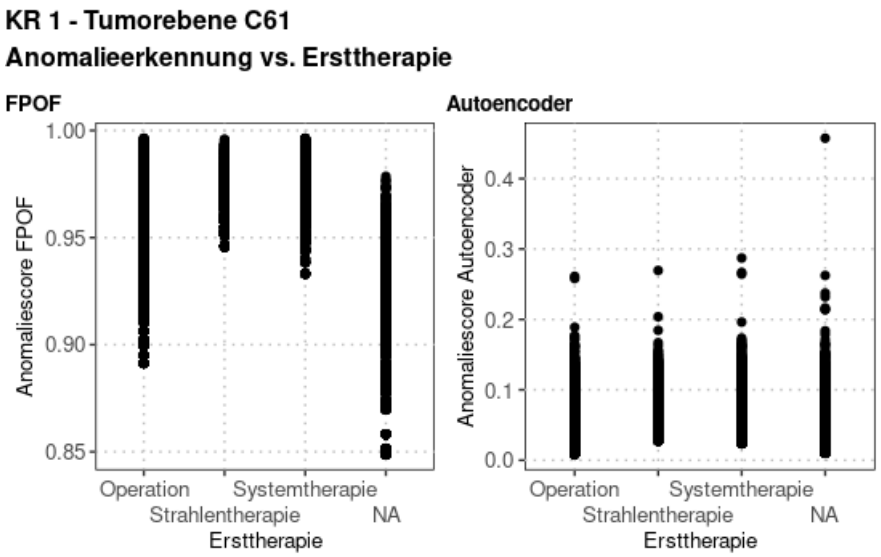
2



3



4



(1) Höherer Anomaliescore mit steigender Anzahl an Therapiemeldungen

(2) Höherer Anomaliescore bei größerem Abstand zwischen Diagnose und Ersttherapie sowie bei negativem Abstand (Therapiedatum vor Diagnosedatum)

(3) Keine Anomalien bzgl. Geschlecht zu erwarten

(4) Keine auffälligen Strukturen bzgl. Ersttherapie (passt zur Leitlinie Prostatakarzinom)

Zweites Untersuchungsfeld – Record Fusion

Ziel:

- Zusammenfügen von Meldungen zu einem Best-Of

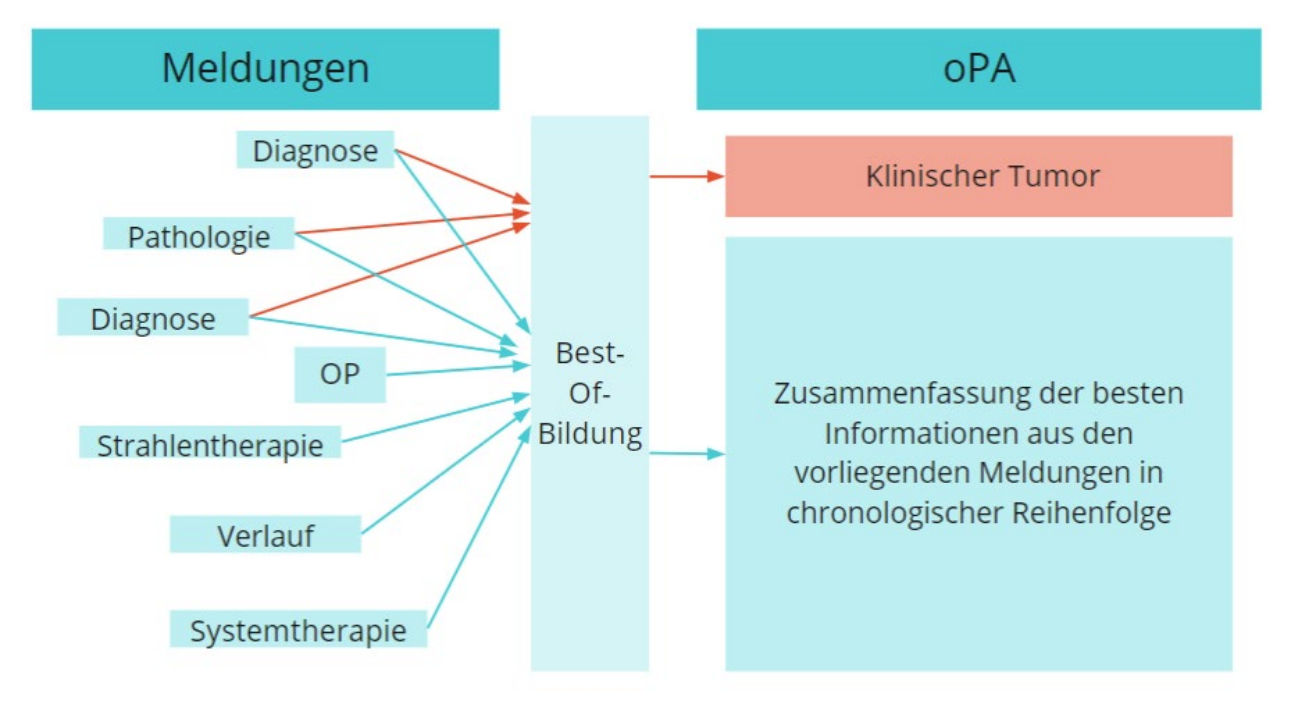
Vorgehen:

- Für jede Variable wurden eigene Modelle trainiert
- Methoden: Supervised Learning (Insb. LSTM)

- **Prädiktoren**: Alle entsprechenden Informationen der Meldungen eines Tumors

- Evaluation: Vergleich der Prediction mit dem Best-Of

- Baseline: Häufigster Wert



Record Fusion

Beispiel mit Variable Seitenlokalisation

Diagnosemeldungen als Prädiktoren

TumorID	Meldung ID	ICD Code	Seiten-lokalisation	Diagnose-sicherung	…
12345	678	C50.4	B	2	
12345	679	C50.4	R	2	
12345	670	C50.4	R	7	

Zielvariable Best-Of

TumorID	Seiten-lokalisation
12345	R

Aus allen Meldungen je Tumor-ID soll ein Best-Of für die Variable vorhergesagt werden

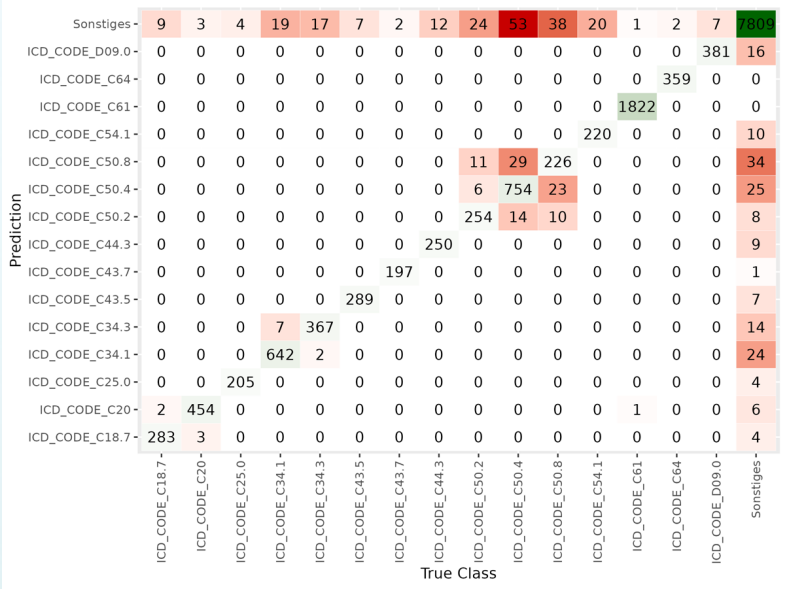
Record Fusion

Übersicht Performance

		KR1		KR2	
Melde-anlass	**Variable**	**Baseline (most frequent)**	**Test accuracy**	**Baseline (most frequent)**	**Test accuracy**
Diagnose	ICD-Code	0.9134	0.9084 (±0.0041)	0.8840	0.8741 (±0.0048)
	ICD-Code (1.-3. Stelle)	0.9754	0.9784 (±0.0019)	0.9714	0.9705 (±0.0157)
	ICDO-Code	0.9156	0.9127 (±0.0029)	0.8516	0.8718 (±0.0043)
	ICDO-Code (1.-3. Stelle)	0.9856	0.9822 (±0.0056)	0.8981	0.9723 (±0.0023)
	Seitenlokalisation	0.9102	0.9393 (±0.0024)	0.8680	0.9556 (±0.0030)
	Seitenlokalisation (nur paarige Tumore)	0.9629	0.9728 (±0.0015)	0.9516	0.9568 (±0.0055)
Histologie	Grading	0.8912	0.9097 (±0.0044)	0.8030	0.9021 (±0.0037)
	Morphologiecode	0.8857	0.9065 (±0.0040)	0.8590	0.9153 (±0.0095)
	Morphologiecode (1.-3. Stelle)	0.9200	0.9413 (±0.0037)	0.9244	0.9502 (±0.0024)
	Morphologiecode (1.-4. Stelle)	0.8852	0.9092 (±0.0051)	0.8914	0.9193 (±0.0056)

→ für den Großteil der Variablen erzeugen die Modelle genauere Ergebnisse als die Baseline
→ Schwierigkeiten gibt es überwiegend bei Variablen mit sehr vielen Ausprägungen
→ Training/Test 70/30, Durchschnittswerte 10 runs cv, 95% Konfidenzintervalle
→ Keine Hyperparameteroptimierung

Ausblick

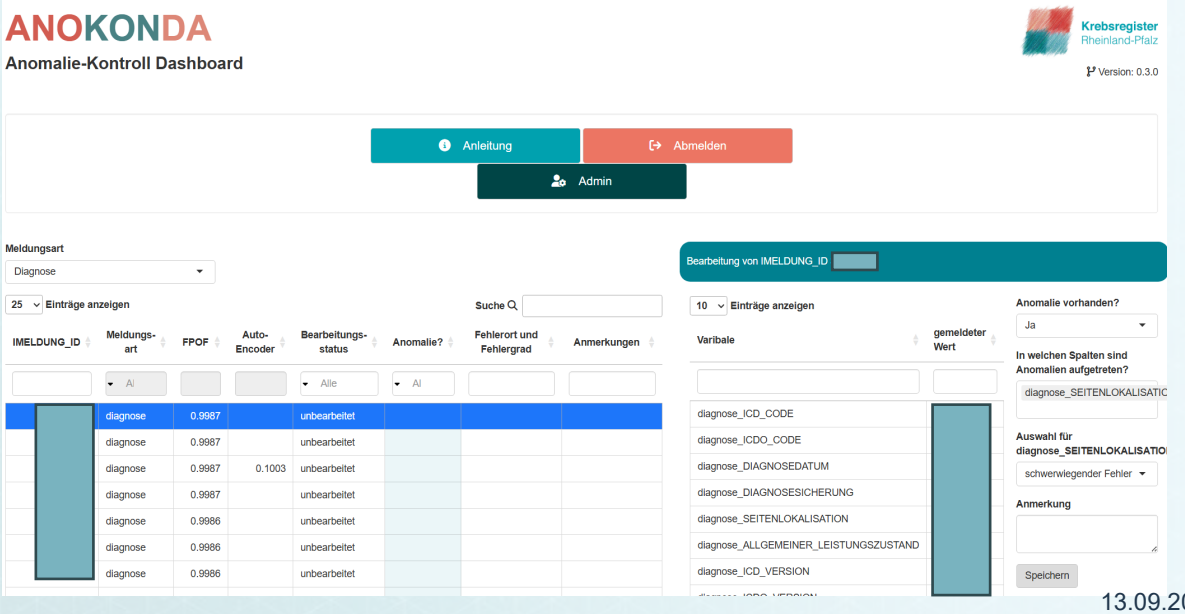


Record Fusion:

- Inwiefern können die Modelle genutzt werden, um in der Praxis Best-Of Werte zu erstellen bzw. vorzuschlagen?
- Modelloptimierung
- Evaluation von Abweichungen

Anomalieerkennung:

- Erstellung eines Anomalie-Dashboards, um direkt Anomalien zu bewerten & zu beheben



Krebsregister
Rheinland-Pfalz
im Institut für digitale Gesundheitsdaten

Für die Gesundheit jedes Einzelnen