

PREDICTIVE ANALYTICS

Was bringt die Zukunft?

DI Dr. Dieter Hayn, MBA

Digital Health Information Systems
Center for Health and Bioresources
AIT Austrian Institute of Technology GmbH
Reininghausstraße 13, 8020 Graz



AGENDA

- Hintergrund
 - AIT & das Projekt HIS-PREMO
 - HIMSS AMAM Skala
- Methoden
 - Übersicht
 - Feature Engineering
 - Modellierung und Evaluierung
- Ergebnisse
 - Wiederaufnahmen (Re-Admission)
 - Abrechnungs-Prüfung durch die Behörde
- Zusammenfassung und Ausblick

AIT Austrian Institute of Technology

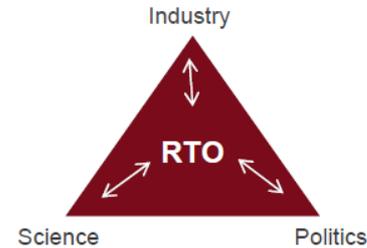
The largest applied research institute in **Austria**



Finance structure



Research and Technology Organisation



Owner structure

50.46%

Republic of Austria

49.54%

Federation of Austrian Industries

1300

Employees

Health & Bioresources	Digital Safety & Security	Vision, Automation & Control
Low-Emission Transport	Technology Experience	Innovation Systems & Policy
Energy	Mobility Systems	

DIGITAL HEALTH INFORMATION SYSTEMS – OFFICE LOCATIONS

~ 40 staff members at 3 locations



Graz



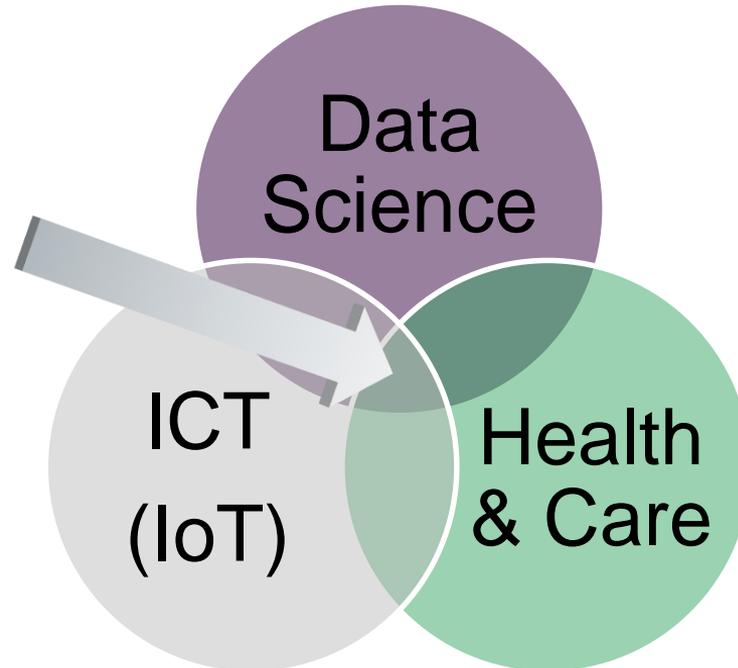
Hall in Tirol
(Innsbruck)



Vienna

PREDICTIVE HEALTHCARE INFORMATION SYSTEMS @ AIT

Information-driven Health & Care



HIS-PREMO

Kooperationsprojekt AIT ↔ Agfa Healthcare

- Einsatz von Methoden des Predictive Modelling
- Im TIP Datawarehouse vorhandene Daten nutzbar machen
- Strukturen und Prozesse im Krankenhaus verbessern

- Aktuell: Wie ist der aktuelle Status?
- Neu: Wie wird der Status in der Zukunft aussehen?

HIS-PREMO, TIP, AMAM SCALE

STAGE	 Adoption Model for Analytics Maturity Cumulative Capabilities
7	Personalized medicine & prescriptive analytics
6	Clinical risk intervention & predictive analytics
5	Enhancing quality of care, population health, and understanding the economics of care
4	Measuring and managing evidence based care, care variability, and waste reduction
3	Efficient, consistent internal and external report production and agility
2	Core data warehouse workout: centralized database with an analytics competency center
1	Foundation building: data aggregation and initial data governance
0	Fragmented point solutions



HIS-PREMO

Re-Admission



HIS-PREMO

Abrechnungs-Prüfung

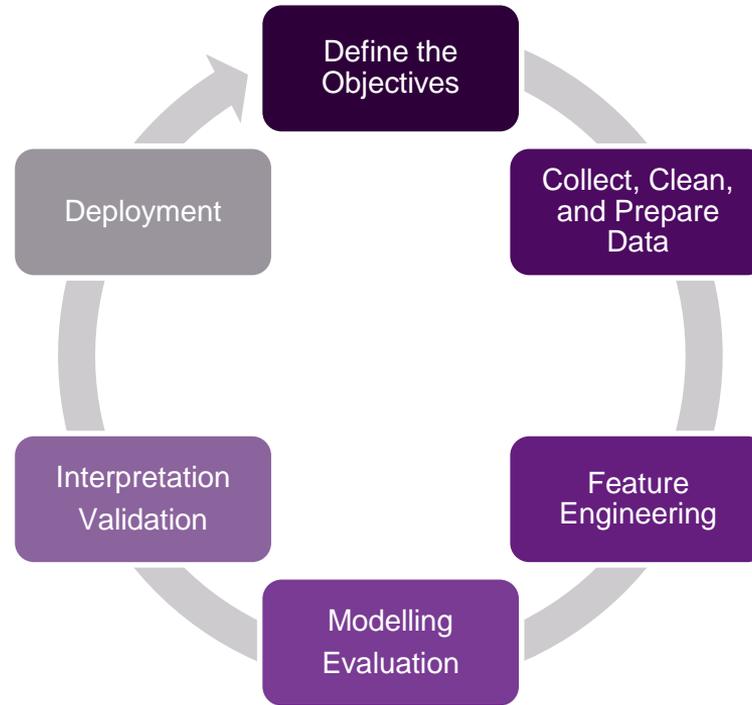
TIP HCe

TIP Solutions

<http://www.himssanalytics.org/amam> Healthcare Information and Management Systems Society

See also EMRAM – Electronic Medical Record Adoption Model

STEPS IN PREDICTIVE ANALYTICS



AUSWAHL ANWENDUNGSFELDER

- Ist die Fragestellung relevant?
- Sind die nötigen Daten vorhanden?
- Sind die nötigen Daten rechtzeitig vorhanden? ...*TIP HCe Live Cubes*...

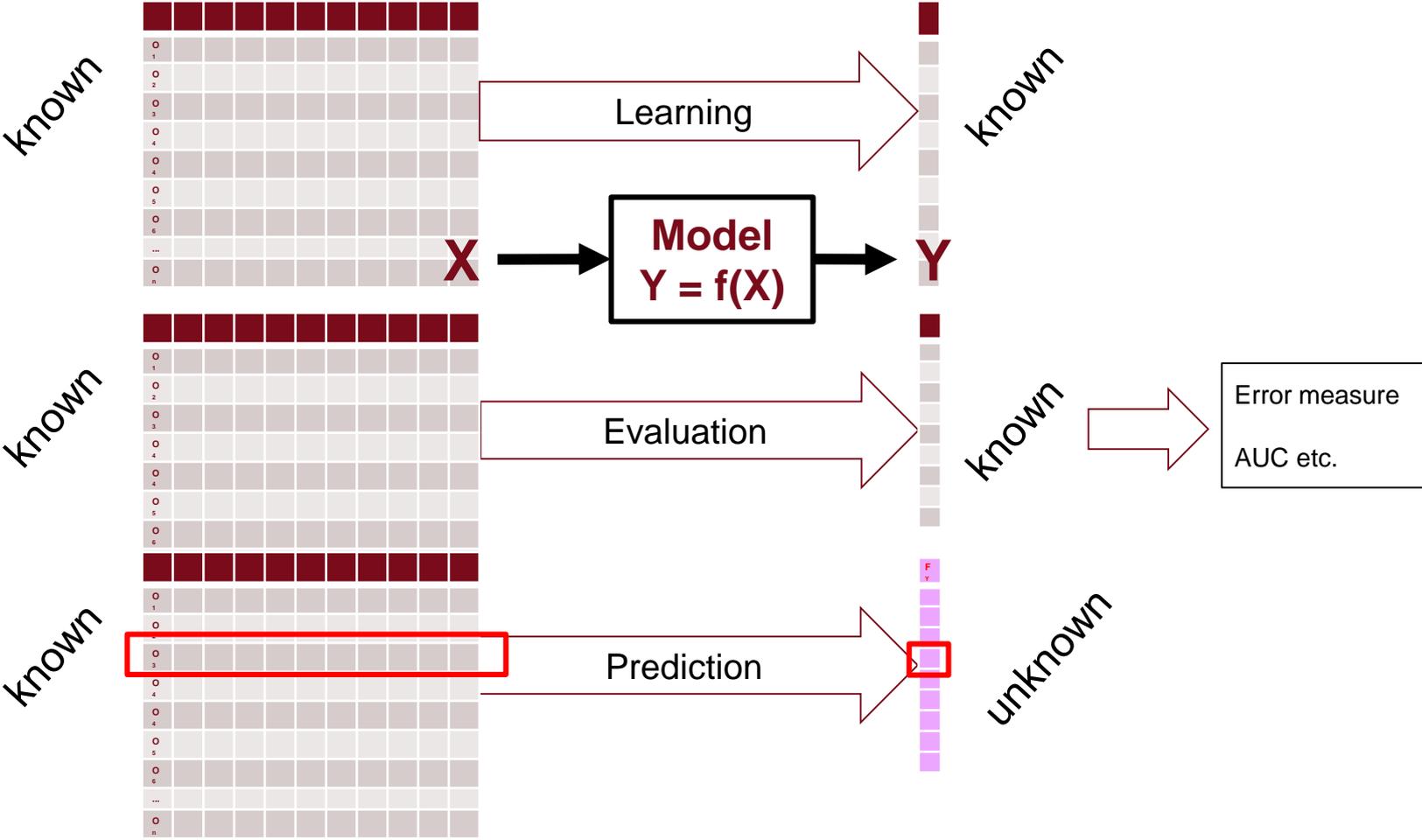
- Aktuell Fokus auf
 - ➔ Vorhersage Fall-Prüfungen durch die Behörde („MDK“, Deutschland)
 - ➔ Vorhersage von Wiederaufnahmen innerhalb von 30 Tagen (Discharge Management)

- Weitere Anwendungsfelder
 - ➔ Vorhersage des Length of Stay / Patient bzw. der Auslastung einer Abteilung
 - ➔ Klinische Fragestellungen (z.B. Vorhersage von Delir bei der Aufnahme)
 - ➔ etc.

DATENQUELLEN

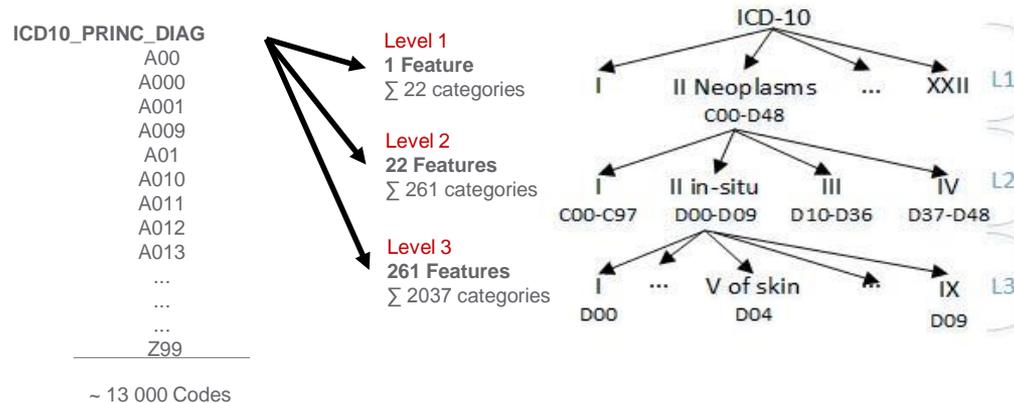
- 1. In allen Krankenhäusern verfügbare, standardisierte Datensätze
 - BFS-Datensatz
 - Abrechnungsdaten (noch nicht vorhanden)
- 2. Spezifische, in einzelnen Krankenhäusern verfügbare Daten
 - KIS / LIS / PACS / PDMS / OP Management System / ...
- Erster Schritt: BFS-Datensätze

PREDICTIVE MODELLING – STAGES



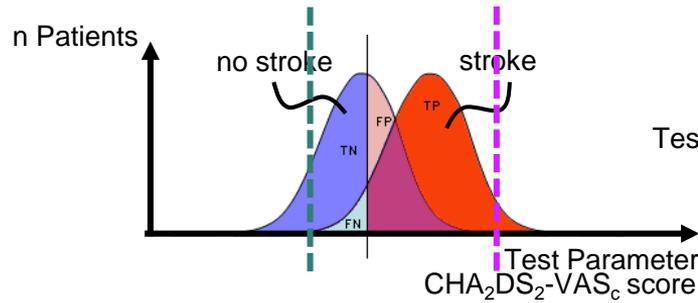
FEATURE ENGINEERING

- Feature Engineering
 - Re-coding
 - Grouping of ICD10 codes

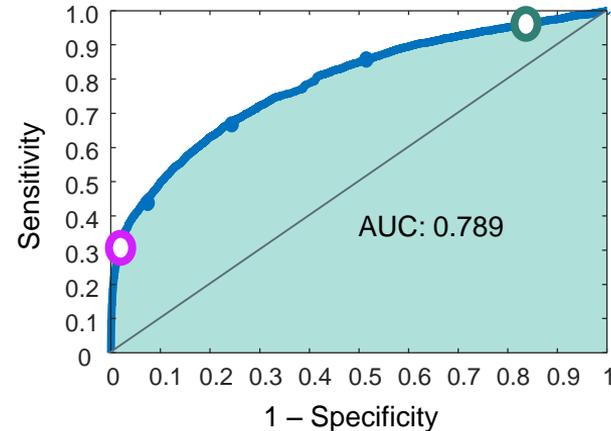


- **Charlson Co-morbidity Index**
 - 1 Feature

RECEIVER OPERATING CHARACTERISTIC (ROC) AREA UNDER THE CURVE (AUC)



		Truth	
		P	N
Test	P	TP	FP
	N	FN	TN



AKTUELLER STAND

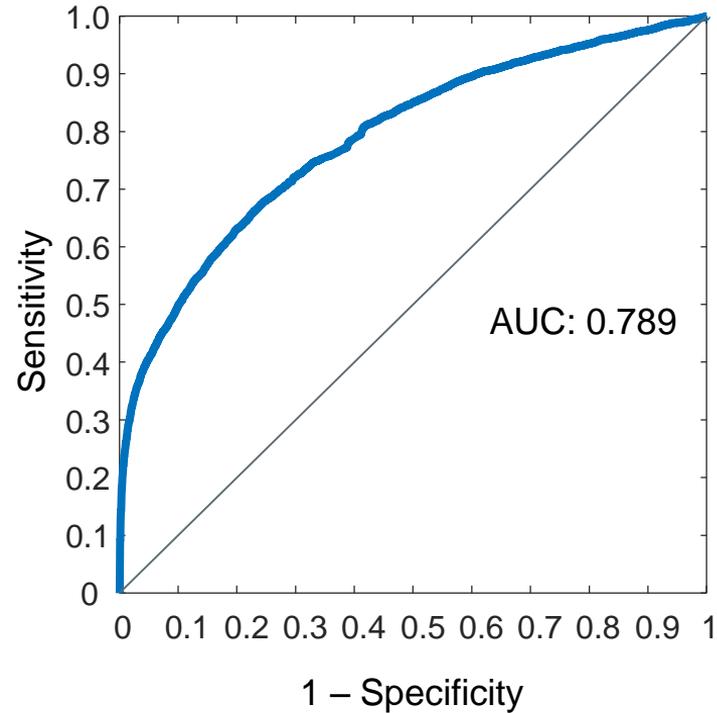
- Daten von mehreren Häusern erhalten
- Feature Engineering
 - Anwendung von Methoden aus vorangegangenen Projekten
 - Spezifische Features für die spezifischen Datensätze
- Initiale Ergebnisse für beide Fragestellungen vorhanden
- Skalierung der Analysen auf mehr Daten rasch möglich

ERGEBNISSE

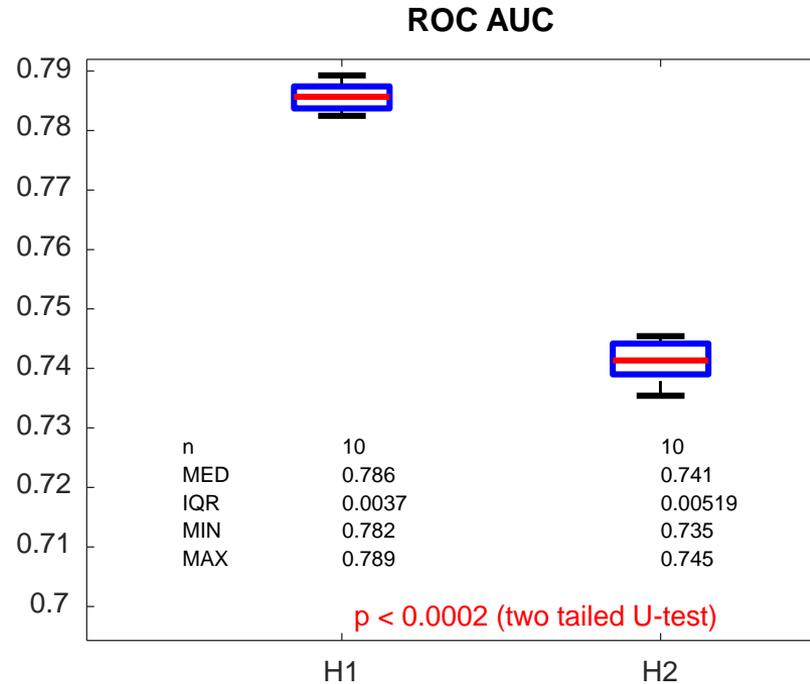
Vorhersage von 30-Tage Wiederaufnahmen

Krankenhäuser		2	
Anzahl Fälle		192.082	Wiederaufnahmen
	KH 1	127.264	17,6%
	KH 2	64.818	7,8%
Wiederaufnahmen		27.430	14,3%

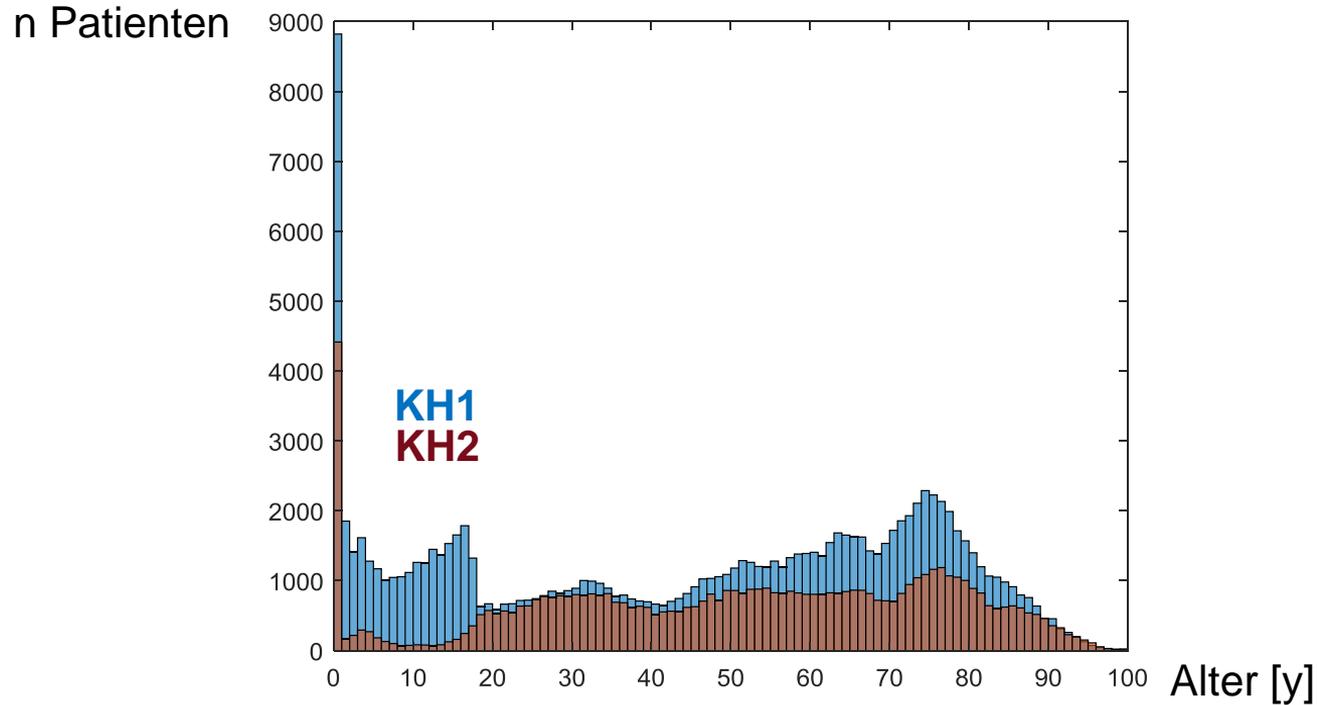
ROC FÜR 30 TAGE WIEDERAUFNAHME VORHERSAGE



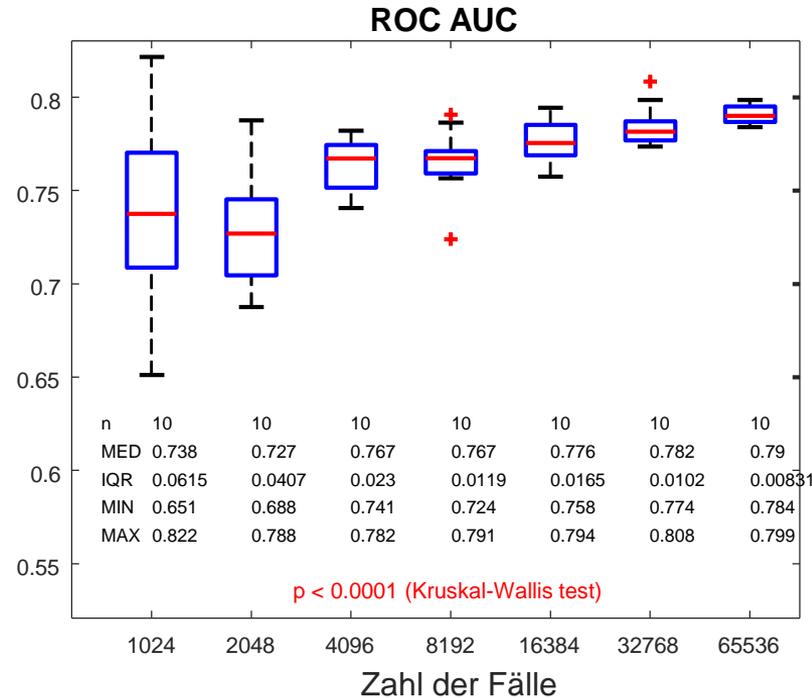
HETEROGENITÄT – ROC GESAMT FÜR HÄUSER 5 UND 6



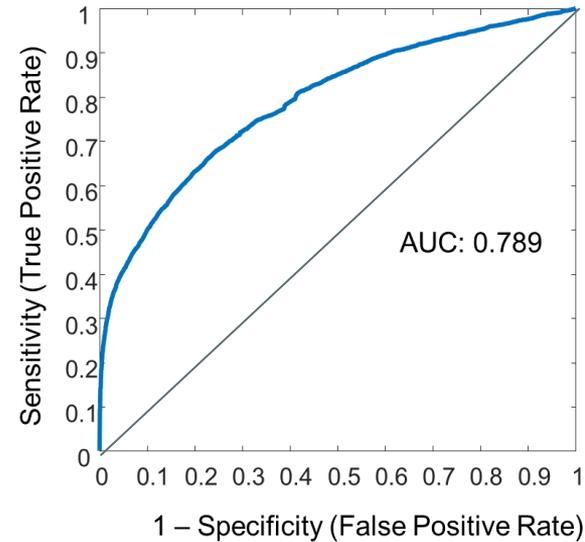
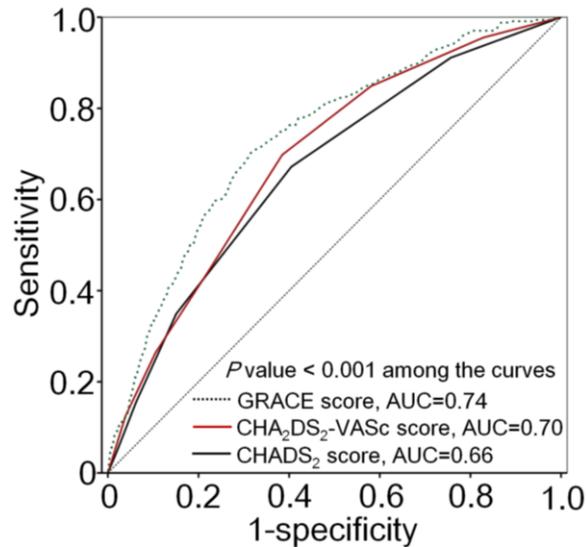
HETEROGENITÄT IN DER PATIENTENDEMOGRAPHIE



EINFLUSS DER DATENSATZGRÖÖE (ZAHL DER FÄLLE)



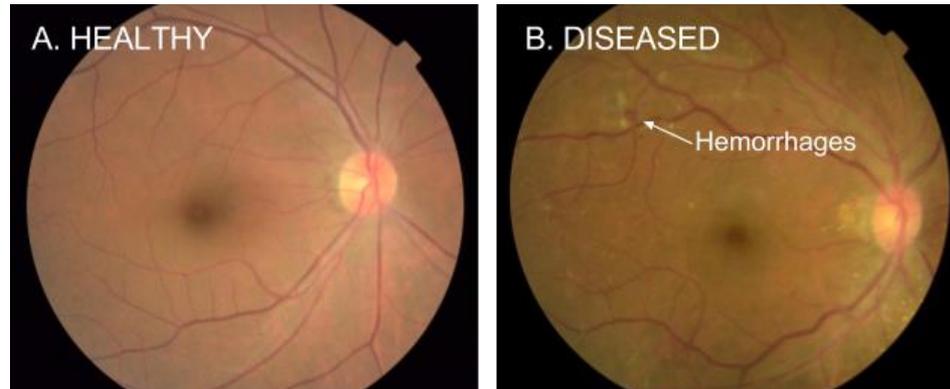
VERGLEICH MIT KLINISCH RELEVANTEN AUC WERTE



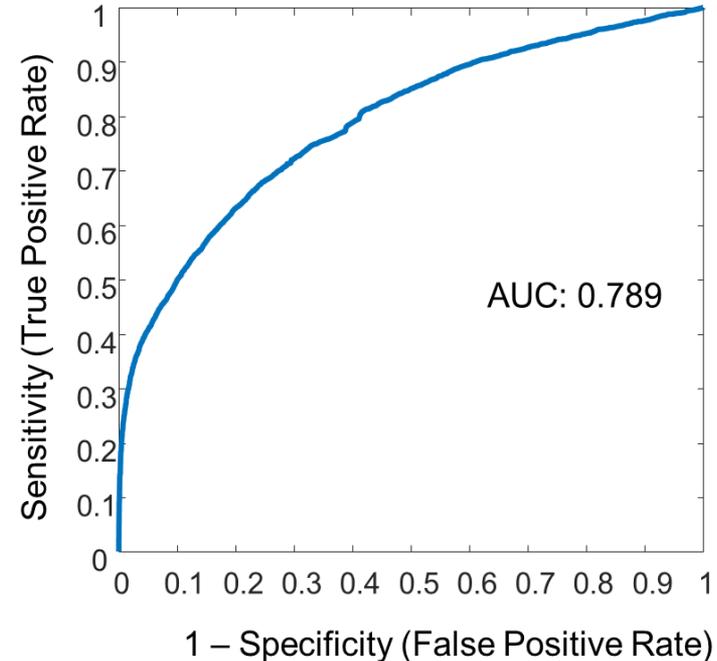
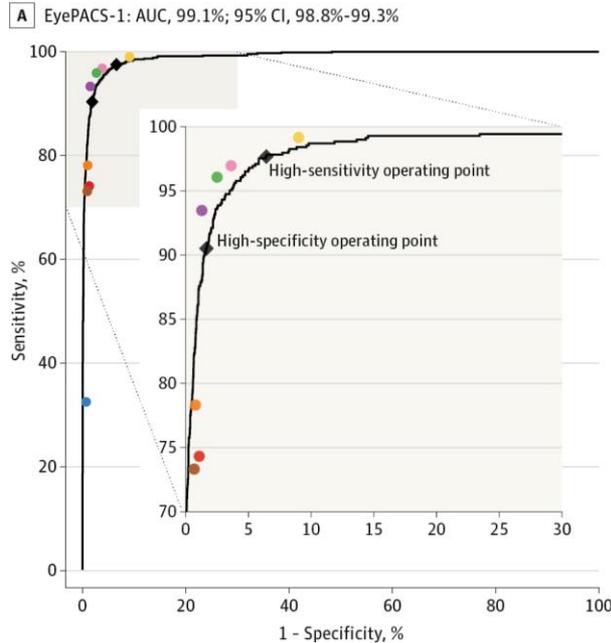
<http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0111167>

EINSATZ VON ARTIFICIAL INTELLIGENCE (AI) BEI DER KLASSIFIKATION VON BILDDATEN

- Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs



VERGLEICH KLINISCH (NOCH) NICHT RELEVANTEN AUC WERTE



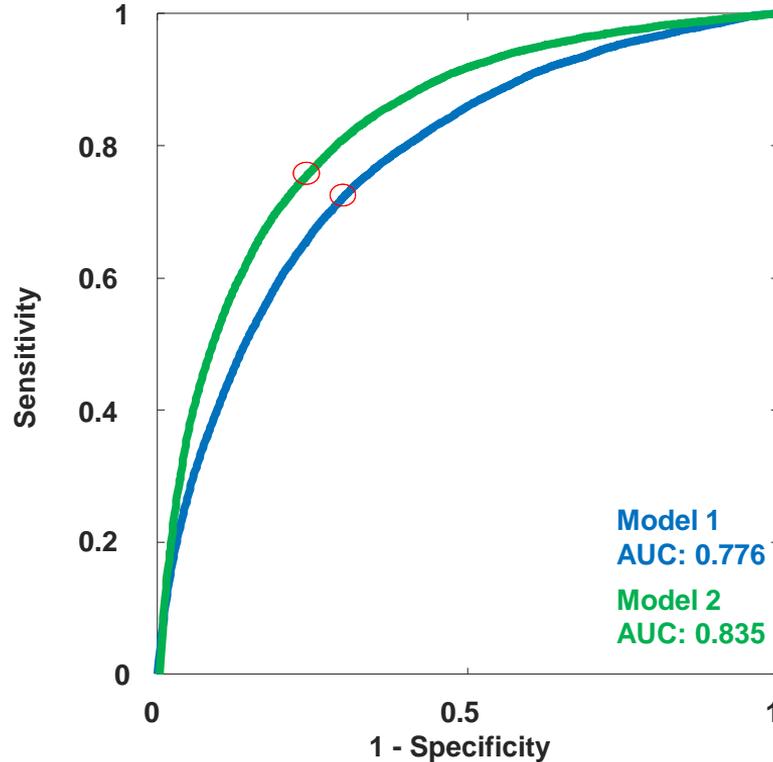
VORHERSAGE VON ABRECHNUNGS-ÜBERPRÜFUNGEN

Anzahl Krankenhäuser	8	
Anzahl Fälle	194.743	
KH 1	74.237	
KH 2	43.910	
KH 3	31.973	
KH 4	15.348	
KH 5	14.711	
KH 6	6.918	
KH 7	5.646	
KH 8	2.000	
Anzahl geprüfte Fälle	26.723	13,7 %

MODELLE UND FEATURES

- Model 1 – 38 Features
- Model 2 – 73 Features
 - 38 Features laut Model 1
 - 35 weitere Features, basierend auf externen Daten
 - Verweildauer in Bezug zur Grenzverweildauer etc. laut Fallpauschalenkatalog
 - Gruppierung von Versicherungen
 - Charlson Comorbidity Index
 - ...

ROC KURVE & KEY-PERFORMANCE-INDIKATOREN BEIDER MODELLE



	Model 1	Model 2
AUC	0,776	0,835
Sens	0,723	0,757
Spec	0,700	0,762

VALIDIERUNG / INTERPRETATION

Wie kommt das Modell (die „Black Box“) auf dieses Ergebnis???

- Feature Importance
 - Wenn alle Werte eines Features **des Lern-Datensatzes** durch Zufallswerte ersetzt würden
→ Wie viel schlechter wäre die Vorhersage?
 - Gibt an, welche Features in der Population viel Einfluss auf die Ausgangsgröße haben
- Individual Feature Influence
 - Gibt an, welche Features **für einen bestimmten Patienten** zur aktuellen Vorhersage geführt haben

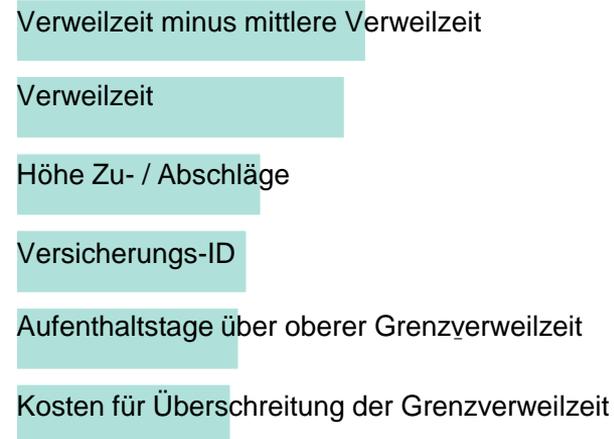
FEATURE IMPORTANCE

Wie kann der **Controller** diese Ergebnisse nutzen?

Die wichtigsten Features in Model 1



Die wichtigsten Features in Model 2



- Anwendung auf gesamten Datenbestand
- Identifikation wesentlicher Einflussfaktoren
- Benchmarking zwischen Abteilungen / Häusern

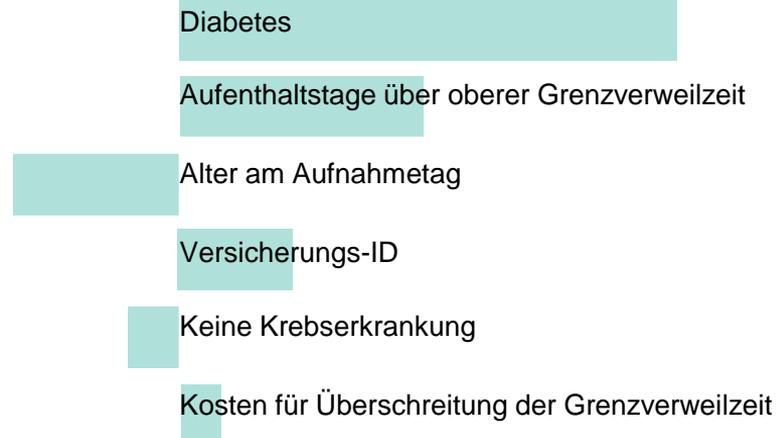
INDIVIDUAL FEATURE INFLUENCE

Wie kann der **Arzt** diese Ergebnisse nutzen?

Globale Feature Importance



Die wichtigsten Features für diesen Patienten



- Anwendung auf die individuellen Daten eines einzigen Patienten
- Plausibilität der Modell-Vorhersage
- Individuelle Entscheidungsunterstützung

ZUSAMMENFASSUNG

- ROC - AUC ~ 0.8 zeigt:
 - In Ihren Daten stecken viele ungenutzte Möglichkeiten
 - In mehr Daten stecken mehr Möglichkeiten
- Herausforderungen
 - Variabilität zwischen Häusern
 - Validierung/Interpretation
 - Anwendungsprozess definieren und implementieren
- Datenspendende Häuser willkommen!

- Was bringt die Zukunft?
 - Predictive Analytics



THANK YOU!

Dieter Hayn, 30.09.2017

