
Potenzielle KI-gestützter Vorhersageverfahren auf Basis von Routinedaten

Zi-Congress Versorgungsforschung

> Versorgung morgen – KI und Krisenresilienz in in der ambulanten Versorgung

>> SESSION: KI-Frühwarnsysteme zur Stärkung der Krisenresilienz

Dr. Thomas G. Grobe (mit Team-Unterstützung), aQua-Institut Göttingen

Berlin, 16.09.2026 (9:15-10:30 Uhr)

AGENDA

- 1 Hintergrund – Abgrenzung des Themas
- 2 Erfahrungen mit Vorhersageverfahren (KI-THRUST)
- 3 Perspektivische Gedanken zu KI und Routinedaten
- 4 Implikationen für Planung, Steuerung, Resilienz



Hintergrund – Abgrenzung des Themas

Routinedaten – hier im engeren Sinne von GKV-Routinedaten

- Wesentliche Inhalte und Strukturen bei allen GKV-Kassen gleich (<https://www.gkv-datenaustausch.de/leistungserbringer/leistungserbringer.jsp>)
- Nahezu alle Daten hochstrukturiert und bislang ohne „Freitexte“ oder Bilder; Diagnosen, (operative) Prozeduren, Arzneimittelsubstanzen in Codes nach einheitlich vorgegebenen Schlüsseln (ICD, OPS, PZN > ATC)
- Der weit überwiegende Teil der Merkmale wird ereignisbezogen erfasst (z.B. Diagnosen nur dann, wenn auch ärztliche Leistung erbracht wurden)
- Daten sind längsschnittlich verfügbar und werden fortlaufend (prospektiv) erfasst
- **Leitfrage: Was ist hier (absehbar) mit KI möglich?**

Erfahrungen mit Vorhersageverfahren ...

Motivation zu KI/ML – Vorerfahrungen mit klassischen Prädiktionsmodellen zu unterschiedlichen, vorrangigen Zielen

1. zur Aufdeckung von statistischen **Zusammenhängen**, auch als Beitrag zur Aufklärung von Wirkmechanismen (z.B. RF von Long-COVID)
2. primär zur Eingrenzung von **Personengruppen** in der Praxis (z.B. besonders gefährdete oder für Maßnahmen geeignete Versicherte ...)
3. zur **Bewertung/Adjustierung** von Kennwerten sowie zur **Vorhersage** von Entwicklungen in Gruppen/Regionen (z.B. Qualitätssicherung, Bedarfsplanung ...)

Übliche Erfahrungen: Vorhersagen können teils brauchbar gelingen,
... zumeist bleibt jedoch auch dann viel Luft nach oben > Bsp. zu (2)



Erfahrungen aus KI-THRUST (G-BA Förderkennzeichen 01VSF20014)

Vorläufer **USER**: Prädiktionsmodelle zur Abschätzung von Risiken für unerwünschte Ereignisse nach Entlassung aus dem Krankenhaus > Unterstützung des Entlassmanagements mit Ampelsystem zu unterschiedlich abgegrenzten Risiken

- Lassen sich Risiken mit innovativen Herangehensweisen aus dem Feld KI/ML nicht doch noch erheblich präziser vorhersagen als mit logistischer Regression???
- ... insbesondere, weil im genannten Beispiel ja eher Vorhersageergebnisse als Einblicke in Zusammenhänge zählen
- Vorbehalte: Ausreißer, nicht nachvollziehbare Ergebnisse, Erklärbarkeit ...

Projektziele: Exemplarische Gegenüberstellung von Ergebnissen unterschiedlicher Vorhersagetechniken, um

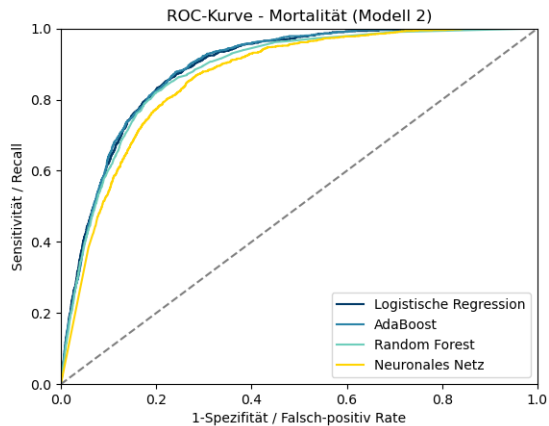
- a) KI-Experten GKV-Routinedaten sowie
- b) GKV-Spezialisten KI-Techniken jeweils näher zu bringen

Ergebnis:

Weißbuch Potenziale KI-gestützter Vorhersageverfahren auf Basis von GKV-Routinedaten

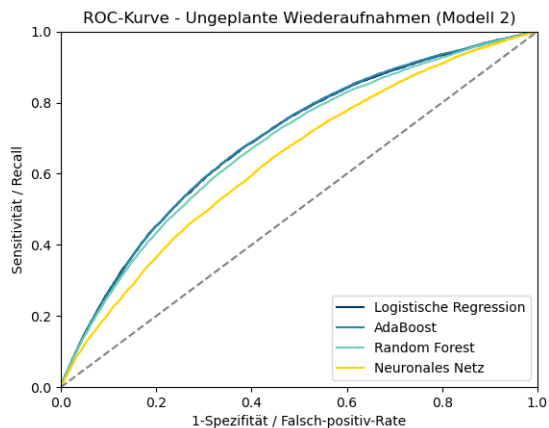


Erfahrungen aus KI-THRUST



Tab. 9-1. Modellgüte Testdaten 2018 basierend auf der Fläche unter der ROC-Kurve (AU-ROC)

Verfahren	AU-ROC
Logistische Regression	0,888
AdaBoost	0,889
Random Forest	0,878
Neuronales Netz	0,861



Tab. 10-1. Modellgüte Testdaten 2018 basierend auf der Fläche unter der ROC-Kurve (AU-ROC)

Verfahren	AU-ROC
Logistische Regression	0,693
AdaBoost	0,694
Random Forest	0,681
Neuronales Netz	0,638

Sterblichkeit nach Entlassung (30d)

- Diagnosen aus dem Jahr vor Aufnahme führen zu deutlich besserer Prognose (Modell 2)
- Die Verfahren **AdaBoost** (ML) und **logistische Regression** bieten die höchste Genauigkeit

Wiederaufnahmen (30d)

- Diagnosen aus dem Jahr vor Aufnahme führen zu leicht besserer Prognose (Modell 2)
- Die Verfahren **AdaBoost** (ML) und **logistische Regression** bieten die höchste Genauigkeit

(Modell 1: Basisprädiktoren Alter, Geschlecht, mehrfache und lange KH-Behandlungen, Polymedikation, Hilfsmittel, Pflegegrad; Modell 2: Basis + 241 Diagnosegruppen; Modell 3: Basis + 4 x 241 Variablen zu Diagnosegruppen differenziert nach Vorquartalen der Krankenhausbehandlung)

Perspektivische Gedanken zu KI und Routinedaten

- Kommunikation nicht einfach: teils unterschiedliche Begriffe für ähnliche Sachverhalte, weder KI/ML-Techniken noch GKV-Routinedaten sind für jeweils Außenstehende inhaltlich einfach verständlich
- Bei identischem Input (Variablen/Features) sind mit alternativen KI/ML-Techniken kaum Vorteile zu erwarten
- Auch Ergebnisse alternativer KI/ML-Techniken hängen entscheidend von der vorausgehenden Aufbereitung ggf. verfügbarer Informationen ab
- KI-Techniken könnten möglicherweise eher im Sinne von Aufbereitung/Vorverarbeitung komplexer Informationen zu Fortschritten führen (z.B. BERT, Word2Vec)
- Ein Umgang mit unstrukturierten Informationen innerhalb von GKV-Routinedaten wäre noch eine komplett andere Herausforderung/weitgehend Neuland

Implikationen für Planung, Steuerung, Resilienz in Krisensituationen und bei Versorgungsengpässen

- Routinedaten liefern wertvolle Informationen für unterschiedliche Kontexte
- Monitoring von Inanspruchnahmen und Erkrankungen/Diagnosen zunehmend etabliert
- In kurzfristigen Krisensituationen schmerzlicher Zeitverzug – mit Ausnahme von AUs
 - > Zeitverzug aktuell deutlich relevanter als ggf. „fehlende KI“
 - > persönlicher Traum: „Echtzeitanalysen“
- KI-Elemente dürften zukünftig auch bei Modellierungen mit hochstrukturierten Routinedaten zu Fortschritten beitragen – am ehesten da, wo klassische Modelle an ihre Grenzen kommen (z.B. Daten mit vielen hochdifferenzierten Merkmalen und exakten zeitlichen Zuordnungen)
- Bei zunehmender Menge an Daten/Merkmalen/Informationen sowie Komplexität von Modellierungen wird es schwieriger, Fehler insbesondere im Detail zu erkennen



Vielen Dank für Ihre Aufmerksamkeit

**aQua – Institut für angewandte Qualitätsförderung
und Forschung im Gesundheitswesen GmbH**

Maschmühlenweg 8–10
37073 Göttingen

Telefon (+49) 0551-789 52-0
Telefax (+49) 0551-789 52-10

office@aqua-institut.de
www.aqua-institut.de

Wir sind ISMS und QMS
zertifiziert nach
ISO/IEC 27001 und ISO 9001

