BLOCKCHAIN® IN HEALTHCARE TODAY
PLATFORM APPROACHES IN HEALTHCARE
TECHNISCHE BRIEFS & KURZBERICHTE

Reduktion der Modellkomplexität für ZKML-Anwendungen im Gesundheitswesen: Schutz der Privatsphäre und Optimierung der Inferenz für ZKML-Anwendungen - eine Referenzimplementierung mit dem synthetischen ICHOM-Datensatz

Sathya Krishnasamy, MS;[1] und Ilangovan Govindarajan, MD[(2)]

[1]Präsident und Direktor, ChainAim, Newington, Connecticut, USA, [2]Chief Medical Officer, GuardianMedx, Las Vegas, Nevada

Korrespondierender Autor: Sathya Krishnasamy, E-Mail: sathya.krishnasamy@chainaim.com

DOI: https://doi.org/10.30953/bhty.v7.340

Stichworte: Blockchain, Diabetes, Distributed Ledger, Modellkomplexitätsreduktion, Datenschutz, maschinelles Lernen, ICHOM, International Consortium for Health Outcomes Measurement, ZKML, Zero-knowledge machine learning

Zusammenfassung

Das Web 3.0 stellt die nächste bedeutende Weiterentwicklung des Internets dar, die die zugrunde liegenden dezentralen Netzwerkarchitekturen, verteilte Ledger und fortschrittliche KI-Funktionen umfasst. Obwohl sich die Technologien rasch weiterentwickeln, gibt es erhebliche Hindernisse für eine weitreichende Einführung. Der Autor erörtert die Hindernisse und die Abhilfemaßnahmen durch spezifische Technologien, die zur Lösung dieser Probleme heranreifen, in einem früheren Papier mit dem Titel Moving Beyond POCs and Pilots, das 2023 in Blockchain in Healthcare Today veröffentlicht wurde. Dazu gehören Technologien zur Wahrung der Privatsphäre, Optimierungen des Off-Chain- und On-Chain-Designs sowie der mehrdimensionale Ansatz, der für die Planung und Einführung dieser Technologien erforderlich ist. Als Erweiterung wird in diesem Beitrag ein solcher Befähiger, Zero Knowledge Machine Learning (ZKML), erörtert, der zwei Technologieströmungen auf einzigartige Weise zusammenführt, um Probleme in Bezug auf den Datenschutz und die Kosten der Inferenz zu lösen. Zero-knowledge proofs (ZKP) ermöglichen es einer Partei, die Gültigkeit einer Aussage gegenüber einer anderen Partei zu beweisen, ohne zusätzliche Informationen über die Aussage selbst preiszugeben. ZKML kombiniert das kryptographische Prinzip von ZKP mit Techniken des maschinellen Lernens (ML). Diese Technologie ist noch nicht ausgereift und benötigt Grundlagen für Anwendungen im globalen Gesundheitswesen. In dieser Arbeit konzipieren die Autoren die technischen und betrieblichen Möglichkeiten der Verwendung von ZKML und implementieren eine Referenzimplementierung im Gesundheitswesen unter Verwendung des synthetischen International Consortium for Health Outcomes Measurement (ICHOM) in der Evaluierungsphase in einer globalen Gesundheitseinrichtung für die Erhebung großer Datenmengen, einschließlich von Patientenberichten über Ergebnisse. Die Reduktion der Modellkomplexität wird für den ICHOM-Diabetes-Datensatz erforscht und berichtet, um die Verwendung von ML-Modellen in globalen Standards der Datenerfassung im Gesundheitswesen in dezentralen Netzwerkarchitekturen für erhöhten Datenschutz und Effizienz zu fördern.

Eingereicht: 1. August 2024; Angenommen: August 25, 2024; Veröffentlicht: August 31, 2024

Die große Menge an Daten, die in den letzten Jahren gesammelt wurde, hat zu noch nie dagewesenen Analysemöglichkeiten geführt. Mit der raschen Zunahme der Datenerfassung und maschinellem Lernen (ML), insbesondere in zentralisierten Systemen, haben die Bedenken hinsichtlich des Datenschutzes und der Sicherheit durch übermäßige Datenbewegungen, die manchmal nicht wirklich erforderlich sind, erheblich zugenommen. Da die zentralen Datenspeicher immer größer werden, werden sie zu attraktiven Zielen. Verstöße gegen den Datenschutz im Gesundheitswesen sind immer häufiger geworden häufig. Es gab viele öffentlichkeitswirksame Vorfälle, und der Trend ist ungebrochen.

Eine wichtige Strategie zur Minderung der mit Gesundheitsdaten verbundenen Risiken ist die Beschränkung des Datenzugriffs an der Quelle und die Reduzierung unnötiger Datenbewegungen. Dazu gehören ein rollenbasierter Zugriff auf die Daten, Datenverschlüsselung, minimaler Datentransfer und regelmäßige Audits. Neben diesen Maßnahmen ist es auch wichtig, Systeme für die Zusammenarbeit zu entwickeln, die über dezentralisierte Netzwerke kommunizieren können. Zero-knowl-edge proofs (ZKP) können für den Nachweis und die Verifizierung zwischen Systemen entwickelt werden, noch vor den Blockchains. Diese

Anhang

ZKPs ermöglichen es einer Partei, die Gültigkeit einer Aussage gegenüber einer anderen Partei zu beweisen, ohne zusätzliche Informationen über die Aussage selbst preiszugeben, und erfreuen sich zunehmender Beliebtheit, da die Distributed-Ledger-Technologie immer ausgereifter wird und zur Skalierung von Blockchain-Implementierungen beigetragen hat.

Das maschinelle Lernen mit Null-Wissen (ZKML) stellt eine revolutionäre Verschmelzung von Kryptographie- und ML-Technologien dar. Es kombiniert das kryptografische Prinzip von ZKP mit ML-Techniken. Durch die Integration von ZKP mit ML stellt ZKML sicher, dass sensible Daten vertraulich bleiben, und ermöglicht dennoch die Entwicklung und Nutzung von Vorhersagemodellen. Diese Integration ist in zunehmendem Maße dort relevant, wo Datenschutz und Sicherheit von größter Bedeutung sind, wie z. B. in Gesundheitssystemen.

In der Praxis ermöglicht ZKML die Zusammenarbeit mehrerer Unternehmen, ohne die Vertraulichkeit ihrer geschützten Informationen zu gefährden. Das bedeutet, dass Organisationen gemeinsam ML-Modelle auf geschützten Datensätzen trainieren und nutzen können, ohne die Daten offenzulegen. So könnte beispielsweise eine medizinische Forschungseinrichtung Daten aus verschiedenen Krankenhäusern zusammenführen, um ein robustes Vorhersagemodell für Krankheiten zu entwickeln, ohne dass ein Krankenhaus seine Patientendaten preisgeben muss. Durch die Verwendung kryptografischer Beweise wird sichergestellt, dass die Daten während des gesamten Prozesses sicher und privat bleiben.

Derzeit befindet sich ZKML noch im Anfangsstadium der Forschung und Entwicklung. Während ZKPs in der kryptografischen Forschung bereits seit den 1980er Jahren ein Thema sind, ist ihre Anwendung auf ML neu und komplex. Diese Technologie steht vor Herausforderungen in Bezug auf Recheneffizienz und Ressourcenbedarf. Die Implementierung von ZKPs kann ressourcenintensiv sein, was die Verarbeitungszeiten und -kosten erhöht.

Die Reduktion der Modellkomplexität ist eine neuere Technik zur Verringerung der Modellkomplexität und damit zur Verkürzung der Berechnungszeiten. Es wurde versucht, einfachere Modelle von Kaggle zu verwenden. Das International Consortium for Health Outcomes Measurement (ICHOM) widmet sich der Entwicklung von Standardsätzen für Ergebnismessungen, die weltweit zur Bewertung der Qualität der Versorgung für verschiedene medizinische Erkrankungen verwendet werden können. Diese Forschung zielt darauf ab, die Verwendung eines globalen Standarddatensatzes für das Gesundheitswesen (ICHOM) zu bewerten und die Modellkomplexitätsreduktion für ein reales, nicht-triviales Beispiel auf einen synthetischen Datensatz im ICHOM-Schema anzuwenden. Daher wird diese Forschung als Referenz für die Entwicklung weiterer Modelle und Optimierungen dienen, die es ermöglichen, die Komplexität und den Einsatz von ZKML für viele Anwendungsfälle im Gesundheitswesen zu optimieren, bei denen Datenschutz und kollaborative Entscheidungsfindung erforderlich sind.

Ein Paradigmenwechsel: Dezentralisierte Systeme, KI-Modelle an der Quelle und kollaborative Entscheidungsfindung

Da wir uns neu entstehenden Architekturen nähern, die auf dezentralisierten Systemen aufbauen, wird die Einschränkung von Daten an der Quelle und das Verständnis von Möglichkeiten, mit den Daten an der Quelle über datenschutzgeschützte ML zu arbeiten, unerlässlich. Daten sind

Es werden wirksame Mechanismen benötigt, um föderiertes Lernen zu ermöglichen, Datenschutz- und Sicherheitsfragen zu lösen und den Rechenaufwand zu verringern. Durch die gemeinsame Nutzung von Modellaktualisierungen anstelle von Daten verbessert das föderierte Lernen den Datenschutz. Auch bei verteilten Ledgern ist ein zunehmend beliebtes Konzept die Bereitstellung von Beweisen für Verifikationssysteme durch ZKP. Dieses Konzept ist zwar nicht neu, doch hat diese Methode Blockchain-Systeme erfolgreich skaliert und wird für verteilte Ledger der nächsten Generation immer beliebter. Diese Mechanismen verringern auch das Risiko der Datenübertragung und stehen im Einklang mit den Datenschutzbestimmungen und der kollaborativen Entscheidungsfindung. ZKML ist eine vielversprechende neue Technologie, die KI/ML-Technologien (künstliche Intelligenz/Maschinelles Lernen) in verteilte Ledger integriert.

Auswirkungen von ZKML auf den Datenschutz im Gesundheitswesen

Daten im Gesundheitswesen sind sensibel; daher ist der Datenschutz das oberste Prinzip bei der Gestaltung von Gesundheitssystemen. Mit ZKML können Gesundheitsdienstleister Erkenntnisse aus ML-Modellen weitergeben, ohne die zugrunde liegenden Patientendaten offenzulegen. Ein Krankenhaus könnte zum Beispiel ein ZKML-Modell zur Vorhersage von Patientenergebnissen verwenden. Das Modell verarbeitet die Daten und erstellt Vorhersagen, während die ZKP sicherstellt, dass diese Vorhersagen korrekt sind, ohne spezifische Patientendaten preiszugeben. ZKML ermöglicht es verschiedenen Einrichtungen, gemeinsam Ergebnisse zu berechnen, ohne ihre individuellen Daten offenzulegen. Dies ist besonders nützlich in der Gesundheitsforschung, wo mehrere Einrichtungen ihre Daten kombinieren möchten, um die Vorhersage von Krankheiten zu verbessern, ohne dabei die Vertraulichkeit der Patientendaten zu gefährden. Da der Datenschutz immer wichtiger wird und der Bedarf an sicheren ML-Modellen wächst, wird die Integration fortschrittlicher Technologien zur Wahrung der Privatsphäre immer wichtiger.

Literaturübersicht

Die ZKPs sind kryptografische Methoden, die es einer Partei (dem Beweisführer) ermöglichen, eine andere Partei (den Überprüfer) davon zu überzeugen, dass eine Aussage wahr ist, ohne zusätzliche Informationen über die Gültigkeit der Aussage selbst hinaus preiszugeben. Goldwasser, Micali und Rackoff (1985)[1]legten den theoretischen Rahmen für interaktive Beweise und ZKPs fest und demonstrierten deren Machbarkeit und grundlegende Bedeutung.

Weitere Arbeiten von Fiat und Shamir[2]dehnten diesen Rahmen auf nicht-interaktive ZKP aus, die nicht mehrere Kommunikationsrunden zwischen Beweiser und Verifizierer erfordern. Jüngste Weiterentwicklungen der ZKP, wie z. B. die Zero-Knowledge Succinct Non-Interactive Arguments of Knowledge (zk-SNARKs) von Ben-Sasson und Kollegen[3,4], haben ihre Effizienz und Skalierbarkeit verbessert und Möglichkeiten für reale Anwendungen eröffnet.

In jüngster Zeit hat ML Erfolge bei der Modellierung von [Vorhersageaufgaben im Gesundheitswesen gezeigt, die von der Krankheitsdiagnose und

Citation: Blockchain im Gesundheitswesen heute 2024, 7: 340 - https://doi.org/10.30953/bhty.v7.340

Prognose zur Patientenbehandlung. Guerra und Kollegen[5]überprüften die ML-Literatur zur Wahrung der Privatsphäre beim Training und bei den Schlussfolgerungen und kamen zu dem Schluss, dass die Datensätze im Gesundheitswesen vielfältig sind und nur ein Bruchteil von ihnen eine Validierung mit unabhängigen Standarddatensätzen in Betracht zieht. Sie wiesen auf die Risiken eines zentralisierten Trainings für föderiertes Lernen hin und betonten die Notwendigkeit einer Zusammenarbeit zwischen verschiedenen Einrichtungen in den verschiedenen Rollen von ML-Wissenschaftlern, Gesundheitspraktikern und Datenschutz- und Sicherheitsexperten, die datenschutzfreundliche Mechanismen für die Zusammenarbeit über verteilte Ledger benötigen.

ZKML ist eine neu aufkommende Technologie, die ZKP auf ML für den Datenschutz anwendet und sicherstellt, dass sensible Daten vertraulich bleiben, während sie die Entwicklung und Nutzung von Vorhersagemodellen für den Datenschutz und die Zusammenarbeit ermöglicht. Die ZKP als solche sind rechenaufwändig und werden für ML-Inferenzbeweise noch rechenaufwändiger. Die jüngste Arbeit von Alejandro Martinez Gator[6]hat eine Bibliothek zur Reduzierung der Modellkomplexität (MCR) entwickelt und ihre Referenzimplementierungen anhand von Musterdatensätzen illustriert. Aus der Perspektive des globalen Gesundheitswesens besteht jedoch die Notwendigkeit, diese Arbeit mit einem standardisierten globalen Gesundheitsschema in großem Maßstab zu validieren und zu vergleichen.

Methoden

Forschungsziele

Guardian Medx ist ein umfassender Pflegeplan, der personalisierte medizinische Versorgung mit kontinuierlicher Überwachung und Unterstützung bietet. Ziel ist es, das Wohlbefinden von Senioren zu verbessern und die Zahl der Krankenhausaufenthalte in Südindien zu verringern. Die Auftraggeber wollten zunächst einen Evaluierungsprozess durchführen, um aus früheren Arbeiten zur Diabetesversorgung in Südindien[7]zu lernen und die kulturellen Elemente zu identifizieren, die in die Krankheitsdynamik involviert sind, von der Diagnose über die Behandlung bis hin zur Adhärenz und kontinuierlichen Pflege, und um ein geeignetes standardisiertes Format zu finden, um sowohl klinische als auch von den Patienten berichtete Ergebnisse auf einer ganzheitlichen Basis zu erfassen. Ziel ist es, ein Schema für die Datenerfassung zu finden, das die Zusammenarbeit unter Wahrung der Privatsphäre in den neuen Web-3-Technologien ermöglicht, einschließlich ML für die Datenanalyse und die Zustimmung und Zusammenarbeit unter Wahrung der Privatsphäre.

Der Ausgangspunkt für diese Forschung ist die Suche nach einem standardisierten globalen Gesundheitsschema, das die oben genannten Herausforderungen bewältigt, eine angemessene Datenerfassung ermöglicht und die für kollaborative Lernzwecke auf globaler Ebene benötigten Konstrukte unter Wahrung der Privatsphäre identifiziert.

Zu den spezifischen Forschungszielen gehören die folgenden:

- Aus früheren Forschungsarbeiten lernen und einen Ansatz zur Datenerfassung entwickeln, der die kulturellen Elemente der Diabetesversorgung erfasst.
- Identifizierung eines spezifischen Datenschemas, das standardisiert ist und für die Datenerfassung und das Lernen unter Wahrung der Privatsphäre auf globaler Ebene verwendet werden kann.
- Untersuchen Sie ZKML als Modell für dieses globale Datenschema.
- Identifizierung der Parameter, Einschränkungen und Abhilfemaßnahmen unter Verwendung von MCR.
- Baseline des Datenschemas mit synthetischen Daten.

Nach Durchsicht der Literatur über die Diabetesversorgung in Indien[7], der großen Datenerhebung und des ML-Aufwands[8]sowie der Berichte über Standardisierungsformate und Anpassungsfähigkeit[9,10]wurde beschlossen, das ICHOM-Schema mit den erforderlichen Anpassungen als Datenschema zu verwenden.

Die Ichom-Datensätze und die Auswirkungen auf die globale Diabetesversorgung

Die ICHOM widmet sich der Entwicklung von Standardsätzen für Ergebnismessungen, die weltweit zur Bewertung der Qualität der Gesundheitsversorgung bei verschiedenen Erkrankungen eingesetzt werden können. Für Diabetes, eine chronische Erkrankung mit erheblichen Unterschieden in der Darstellung und Behandlung in verschiedenen Kulturen und Gesundheitssystemen, ist diese Standardisierung von entscheidender Bedeutung.

Die Diabetesversorgung kann sich aufgrund kultureller Unterschiede, sozioökonomischer Faktoren und der Infrastruktur des Gesundheitswesens erheblich unterscheiden. So sind beispielsweise die Behandlungsstrategien und Patientenergebnisse in einem Land mit hohem Einkommen möglicherweise nicht direkt auf ein Umfeld mit niedrigem Einkommen und unterschiedlichen Ressourcen und kulturellen Einstellungen zur Gesundheit anwendbar. Der ICHOM-Datensatz geht auf diese Herausforderung ein, indem er es den Gesundheitssystemen ermöglicht, standardisierte Maßnahmen an lokale Gegebenheiten anzupassen. Durch diese kulturelle Anpassung wird sichergestellt, dass die Ergebnismessungen in verschiedenen Umfeldern relevant und praktikabel sind, wodurch der Nutzen des Datensatzes und die Auswirkungen auf die globalen Gesundheitsergebnisse verbessert werden.

Durch die Bereitstellung eines Standardsatzes von Messgrößen hilft ICHOM, Lücken in der Versorgung und bei den Ergebnissen in verschiedenen Regionen und Bevölkerungsgruppen zu erkennen. Durch einen einheitlichen Ansatz bei der Messung von Ergebnissen wie Blutzuckerspiegeln, Lebensqualität und Komplikationsraten ermöglicht der ICHOM-Diabetes-Datensatz Gesundheitsdienstleistern, ihre Leistungen mit globalen Standards zu vergleichen, bewährte Praktiken zu ermitteln und die Patientenversorgung und die Gesundheit der Bevölkerung zu verbessern. Angesichts der Beschaffenheit der Daten und der Bemühungen um die Standardisierung begannen die Auftraggeber die Bewertung mit synthetischen Datensätzen. Sie verwendeten die ICHOM-Datensätze für die ältere Bevölkerung und für Diabetiker, um Einblicke in die Komplexität der Diabetesversorgung und ihre Auswirkungen auf die Patientenergebnisse zu gewinnen.

Diese Forschung bietet einen umfassenden datenorientierten Ansatz für das Diabetesmanagement, indem sie Daten zu Demografie, Diagnose, Lebensstil und sozialen Faktoren, Behandlungsmethoden, Diabeteskontrolle, akuten Ereignissen, chronischen Komplikationen und von Patienten gemeldeten Ergebnissen sammelt und analysiert.

Ergebnisse. Ziel ist es, das komplexe Zusammenspiel der Faktoren, die sich auf das Diabetesmanagement auswirken, zu erhellen und die Diabetesversorgung kosteneffizient zu optimieren, indem eine frühzeitige Diagnose sowie eine kontinuierliche Fernüberwachung in großem Maßstab ermöglicht wird.

Als Teil der Datenauswertung für die Forschung werden synthetische Datensätze mit hypothetischen Werten aus der früheren Literatur und den Erfahrungen der Forscher erstellt. Für den praktischen Einsatz im indischen Kulturkreis werden Datenadäquanz, Basisdaten und aussagekräftige Datenmappings mit den ICHOM-Diabetes-Datensätzen V5.0 erstellt. Diese Datensätze wurden aus dem ICHOM V5-Diabetes-Datensatz ausgewählt. Synthetische Daten wurden für 100 Patienten erstellt, wobei mehrere Iterationen der Daten und klinische Validierungen durchgeführt wurden. Es wurde eine explorative Datenanalyse mit univariaten und bivariaten Analysen und Korrelationen entwickelt und ausgewertet.

Das Modell wurde mit einem Light-Gradient-Boosting-Machine (LightGBM)-Regressor erstellt - einem Open-Source-Hochleistungs-Gradient-Boosting-Framework, das für effiziente und skalierbare ML-Aufgaben entwickelt wurde. Es wurde speziell für Geschwindigkeit und Genauigkeit entwickelt und ist daher eine beliebte Wahl für strukturierte und unstrukturierte Daten in verschiedenen Bereichen. Zu den Hauptmerkmalen von LightGBM gehören die Fähigkeit, große Datensätze mit Millionen von Zeilen und Spalten zu verarbeiten, die Unterstützung für paralleles und verteiltes Rechnen sowie optimierte Gradient-Boosting-Algorithmen, die Histogramm-basierte Techniken und blattweises Baumwachstum verwenden.

Ein entscheidender Aspekt von ZKML ist die Reduzierung der Modellkomplexität, die angesichts der aktuellen Inferenzkosten und der Skalierbarkeit der Distributed-Ledger-Technologie entscheidend ist, um diese Modelle effizienter und praktikabler für reale Anwendungen zu machen. In diesem Artikel wird das Konzept der Modellkomplexitätsreduzierung im Zusammenhang mit ZKML anhand von konkreten Beispielen untersucht, wobei der Schwerpunkt auf der entscheidenden Rolle im Gesundheitswesen liegt. Die Reduktion der Modellkomplexität beruht auf den Konzepten des "Pruning", d. h. des Entfernens unnötiger Teile des Modells, die nur minimal zur endgültigen Entscheidung beitragen, der Quantisierung, d. h. der Verringerung der Genauigkeit von Gewichten und Aktivierungen für eine effiziente Berechnung, und der Wissensdestillation, d. h. der Übertragung des Wissens in ein einfacheres Modell, das dennoch Vorhersagefähigkeiten besitzt. Dieses reduzierte Modell kann dann in einem ZKML-Rahmen verwendet werden, um Berechnungen effizient durchzuführen und gleichzeitig die Privatsphäre durch ZKPs zu gewährleisten. Ein weiteres Ziel war es, die technische Durchführbarkeit der für ZKML-Anwendungen generierten synthetischen Daten in Beweis- und Verifikationssystemen zu ermitteln, da das Potenzial für Cross-Learning-Einsichten gegeben ist.

ohne Verlust der Privatsphäre.

Reduktion der Modellkomplexität für den synthetischen ICHOM-Diabetes-Datensatz

Die Reduzierung der Modellkomplexität in ZKML ist wichtig, um Überanpassung zu vermeiden, die Interpretierbarkeit zu verbessern und die Recheneffizienz zu erhöhen, indem die Rechenressourcen für Training und Inferenz reduziert werden. Die ZKML-Software und ihre Bibliothek zur Modellreduktion, die für diese Forschung verwendet wurde, war GIZA ZK Cook.

Der Komplexitätsreduktionsalgorithmus führt die folgenden Schritte aus.

1. Korrelationsanalyse und Merkmalsbedeutung für die Merkmalsauswahl und -reduktion: Merkmale mit hoher Korrelation sind Kandidaten, die zur Redundanz und Reduktion beitragen können. Mit Hilfe von Techniken wie der rekursiven Merkmalseliminierung werden Merkmale mit geringer Wichtigkeit eliminiert.
2. Die L1 (Lasso)-Regularisierung setzt weniger wichtige Merkmalskoeffizienten auf Null, und die L2 (Ridge)-Regularisierung bestraft große Koeffizienten, um die Komplexität zu verringern, ohne Merkmale zu eliminieren. Die L1- und L2-Regularisierung werden kombiniert, um die Vorteile beider Verfahren auszugleichen. Bei baumbasierten Modellen werden Zweige, die nur minimal zu den Vorhersagen beitragen, durch Beschneidung entfernt und Knoten konsolidiert und geteilt.
3. Die Hauptkomponentenanalyse (PCA) und t-distributed Stochastic Neighbor Embedding (t-SNE) werden zur Dimensionalitätsreduktion eingesetzt.
4. Multi-Pass-Kreuzvalidierung und Abstimmung der Hyperparameter zum Ausgleich von Modellkomplexität und Genauigkeit.

Das MCR wurde aus der Giza-Bibliothek verwendet. Das Analysemodell läuft in einer Python-Umgebung.

Ergebnisse und Diskussion

Die Datenkalibrierung aus der Evaluierungsphase zeigt, dass die ICHOM v5-Datensätze in der regelmäßigen und kontinuierlichen Überwachung zur Verhinderung des Fortschreitens von Erkrankungen und zur Verringerung der Lebensqualität (Abbildung 1) praktisch und in erheblichem Umfang genutzt werden können, wie sie mit den Lebensqualitäts-Scores (Abbildung 2) in Beziehung stehen und wie die Adhärenz granular abgebildet werden kann (Abbildung 3).

Die klinische Bewertung von Guardian Medx von synthetischen Daten, die auf der Grundlage einer anonymisierten Extraktion früherer aggregierter Ergebnisse für die südindische Bevölkerung erstellt wurden, hat ergeben, dass das Modell wesentlich dazu beitragen kann, ergebnisorientierte und kosteneffiziente Fernbehandlungen voranzutreiben.

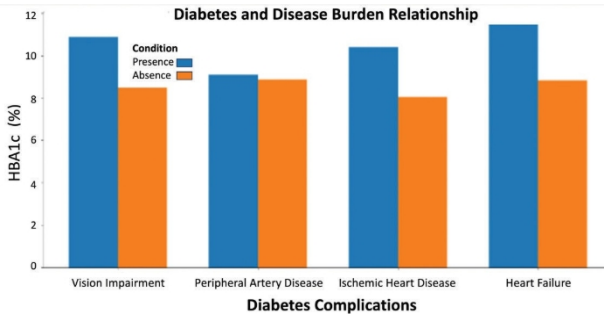
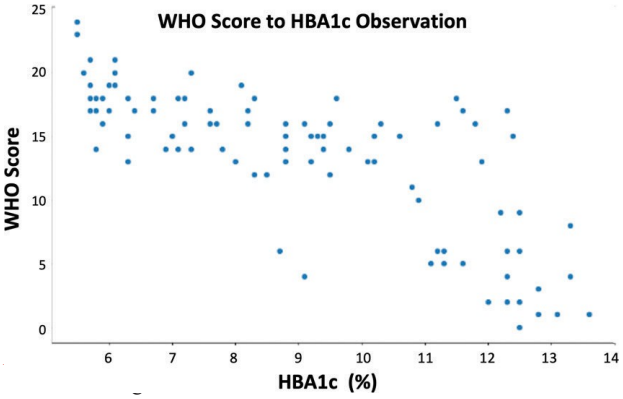


Abbildung 1. Komplikationen bei Diabetes, die die Lebensqualität beeinträchtigen. HBA1c: glykiertes Hämoglobin.
Quelle: Copyright der Autoren

Citation: Blockchain im Gesundheitswesen heute 2024, **7**: 340 - https://doi.org/10.30953/bhty.v7.340

Überwachung, basierend auf der Datenanalyse und den Regressor-Modellen für die in Phase 1 analysierten Daten. Dies ist besonders



(Weltgesundheitsorganisation) im Vergleich zum HBA1c (glykiertes Hämoglobin).
Quelle: Copyright der Autoren

Dies gilt insbesondere für die Kontrolle des HBA1c (glykiertes Hämoglobin), die Häufigkeit der Nachverfolgung und die Beziehung zu den Komplikationen und den Lebensqualitätswerten. Die Koexistenz anderer chronischer Krankheiten war aus den Daten ersichtlich. Die ICHOM-Daten erfassen die von den Patienten angegebenen Ergebnisse mit WHO-Scores, die einen negativen Trend bei den Lebensqualitäts-Scores mit steigenden HBA1C-Werten zeigten. Alle Metriken zur Therapietreue zeigten ebenfalls den erwarteten Zusammenhang mit den HBA1c-Ergebnissen.

Die Studie eröffnet die Möglichkeit, auf der Grundlage von Eskalationsprognosen Maßnahmen zu ergreifen, wenn die Fernüberwachungsdaten für das Patientenmanagement im Hinblick auf frühzeitige Interventionen zur Verfügung stehen.

Modellkomplexität vor der Modellreduktion

Der Hauptaspekt ist die Bewertung der Modellkomplexitätsreduzierungs-Baselines für diesen synthetischen Datensatz und der Vergleich der

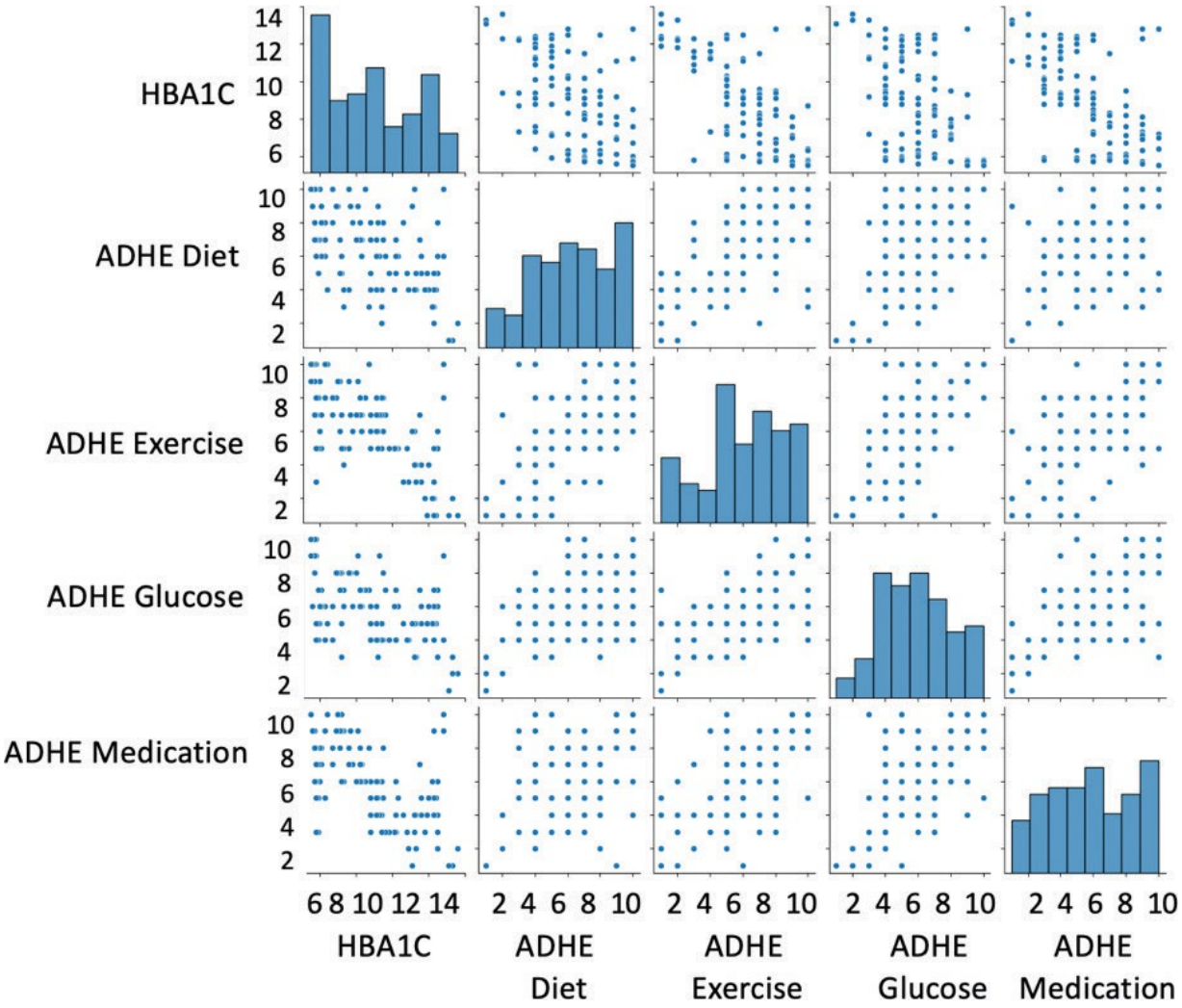


Abbildung 3. Auswirkung der Adhärenz (ADHE) auf die Kontrolle des HBA1c (glykiertes Hämoglobin). Quelle: Copyright der Autoren

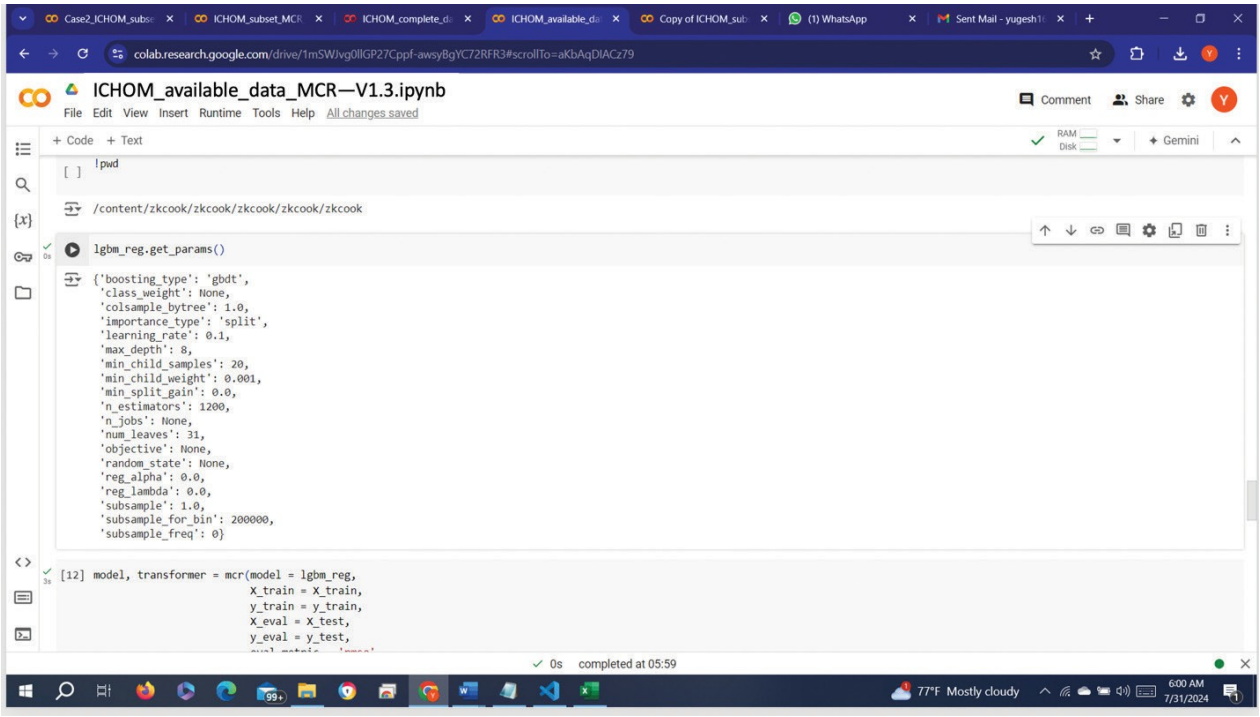


Abbildung 4 Modellkomplexität vor Reduzierung der Modellkomplexität. ICHOM: International Consortium for Health Outcomes Measurement; MCR: Modellkomplexitätsreduzierer.
Quelle: Copyright bei den Autoren

die Komplexität der Modellparameter vor und nach der Reduktion. Dies ist von entscheidender Bedeutung für die Optimierung der Laufzeitkosten und um zu sehen, wie gut sich dieses Modell für den Einsatz in verteilten ML- und kollaborativen Systemen über verteilte Ledger eignet.

Der LGBM-Regressor wurde mit den Parametern n-estimators: 1.200 und max-depth: 8 (Abbildung 4). Im Gegensatz dazu wird die Modellkomplexität auf n-estimators: 150 und max-depth: 4, nachdem die Modellkomplexität nach dem Durchlaufen der ZKCook-Bibliothek reduziert wurde (Abbildung 5).

Dies wird in Form von Knotenpunkten dargestellt:

- Anzahl der Knoten= Anzahl der Bäume* (2 $^{\text{Tiefe}}$-1)
- Die Anzahl der Knoten vor der Komplexitätsreduzierung beträgt 1.200 * (2 $^{(8)-1}$) = 306.000
- Die Anzahl der Knoten vor der Komplexitätsreduzierung beträgt 150 * (2 $^{(4)-1}$) = 2.250

Modellkomplexität nach Reduktion der Modellkomplexität

Die Differenz entspricht einer Reduktion von 99,26 % und steht auch im Einklang mit einigen anderen Referenzbeispielen. Die Wahl des Regressors für das Problem und die weitere Reduzierung mit MCR auf der Grundlage der Evaluierungsdaten zeigt, dass der ICHOM V5 Diabetes-Datensatz zur Erfassung von Daten verwendet werden kann, so dass die Interoperabilität verbessert werden kann, um Ergebnisse in kollaborativen Einstellungen zu untersuchen und zu melden, die in ZKML-Anwendungen verwendet werden können.

Diese Ergebnisse deuten darauf hin, dass die für die Diabetes-Studie benötigte Anpassungsfähigkeit im globalen For-mat von ICHOM erfasst werden könnte, und sind auf der Grundlage der Evaluierungsdaten vielversprechend. Darüber hinaus sind diese Ergebnisse vielversprechend bei der Generierung von datenschutzgeschützten Beweisen für Verifikationssysteme und bei der gemeinsamen Nutzung von Diagnosenachweisen, die auf der Zustimmung des Patienten beruhen, auf datenschutzgeschützte Weise mit anderen kooperierenden Parteien. Angesichts dieser Ergebnisse in der Evaluierungsphase können die Beweise nach Fortschreiten der Studie mit optimierter Rechenleistung erstellt werden.

Diese spezifische Anpassung umfasste eine Teilmenge des vollständigen Wörterbuchs des ICHOM-Datensatzes, da dieser an dieses klinische Umfeld angepasst wurde, was eine Einschränkung darstellt. Wir sehen dies als Ausgangspunkt für weitere Arbeiten zur Verwendung von standardisierten Datensätzen wie den ICHOM-Datensätzen in verschiedenen Kohorten und anderen Krankheiten aus den ICHOM-Datensätzen. Einige dieser Situationen werden eine größere Anzahl von Datenspalten und Anforderungen an die Datenanalyse haben, was uns zusätzliche Referenzpunkte für die Komplexität vor und nach den Baselines und deren Auswirkungen auf Nachweis- und Verifizierungssysteme geben wird. Außerdem ist zu beachten, dass es sich hierbei um eine vorläufige Basislinie handelt, da die Technologie in allen Bereichen - Entwicklung von Standards, Modellreduzierungstechniken, Reduzierung von Beweisen und Beschleunigung von Verifikationssystemen sowohl auf der Software- als auch auf der Hardwareebene - sehr schnell reift. Daher wird es sich

Citation: Blockchain im Gesundheitswesen heute 2024, **7**: 340 - https://doi.org/10.30953/bhty.v7.340

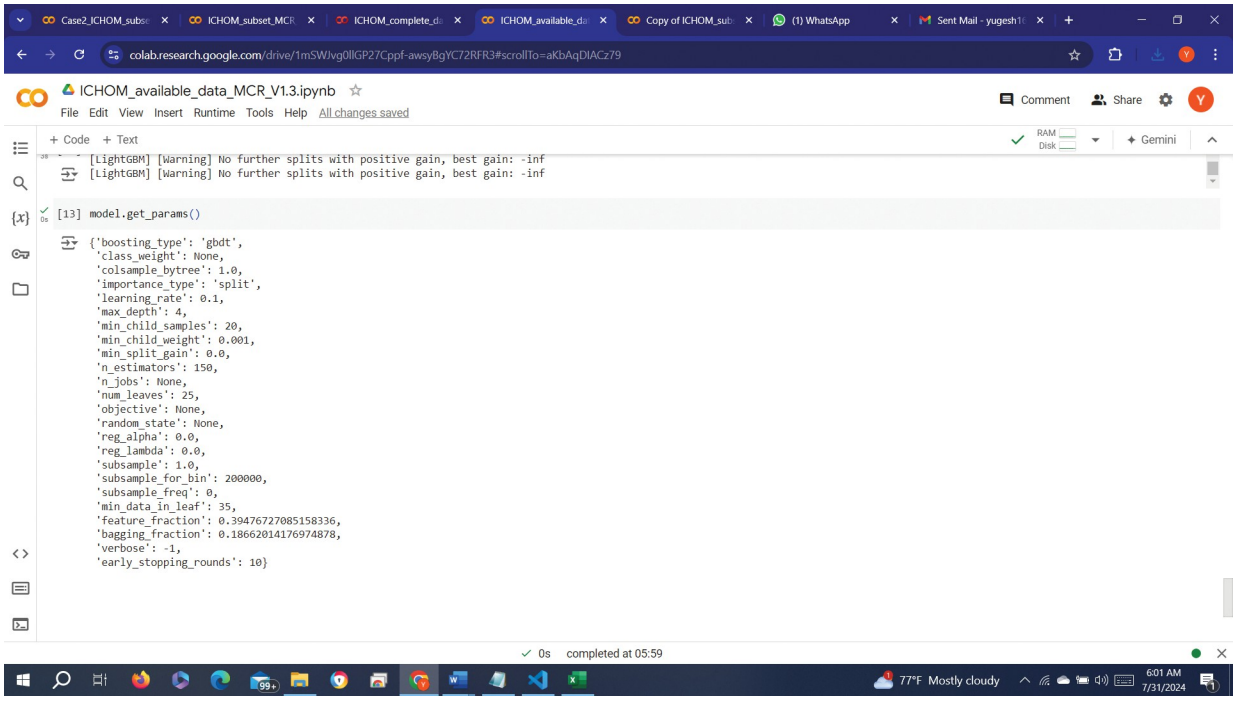


Abbildung 5. Modellkomplexität vor der Reduzierung der Modellkomplexität. ICHOM: International Consortium for Health Outcomes Measurement; MCR: Modellkomplexitätsreduzierer.
Quelle: Copyright by the authors

Es ist wichtig, ein Register der ZKML-Entwicklungen über diese Parameter hinweg zu führen.

Schlussfolgerungen und zukünftige Arbeiten

Auf der Grundlage der Datenanalyse und der Regressormodelle für die in der Evaluierungsphase analysierten Daten ist die klinische Evaluierung von GuardianMedx der Ansicht, dass das Modell die ergebnisorientierte und kosteneffiziente Fernüberwachungsversorgung erheblich voranbringt. Die Anwendung des Giza ZKcook Algorithmus zur Reduktion der Modellkomplexität auf ICHOM-Diabetesdaten führte zu besser interpretierbaren und rechnerisch effizienteren Modellen. Die Rechenzeiten wurden auf einem standardisierten ICHOM-Datensatz erheblich reduziert, um ML- und datenschutzgeschützte Einstellungen zu verwenden und die Daten an der Quelle zu halten. Dies erhöht die Sicherheit und liefert überprüfbare Beweise für alle Vorhersagemodelle, die Agenten steuern, und die Verwendung von ML-Modellen in Verbindung mit dezentralisierten verteilten Leitern, um Möglichkeiten der Zusammenarbeit zu eröffnen, ohne die internen Details der Daten preiszugeben. Angesichts dieser Ergebnisse in der Evaluierungsphase können die Beweise nach Fortschreiten der Studie mit optimierter Rechenleistung erstellt werden. Weitere Arbeiten können auf andere Anpassungen des ICHOM-Frameworks für Diabetes in einer anderen Kohorte in einer anderen Umgebung ausgedehnt werden, um die Ergebnisse zu vergleichen, sowie auf andere Krankheitsdatensätze von ICHOM. In Kürze plant das Team, die ZKML-Funktionalitäten zu erweitern, um Agenten für die weitere Verarbeitung zu füttern und geschützte Mehrparteienerkenntnisse voranzutreiben.

Finanzierung

Keine.

Interessenkonflikte

Keine.

Mitwirkende

Sathya Krishnasamy ist der Präsident und Leiter von ChainAim Technologies. Er verfügt über 25 Jahre Erfahrung im Bereich Managed Care bei führenden US-Gesundheitsunternehmen, darunter Aetna und Anthem. Er konzentriert sich auf aufkommende Technologien, einschließlich KI/ML-Systeme und Distributed-Ledger-Technologien. Darüber hinaus ist er als Berater bei zahlreichen Bemühungen der Branche im Bereich der Zusammenarbeit zwischen Kostenträgern und Leistungserbringern, Standardisierungsorganisationen und Bemühungen wie Account Aggregators in Indien tätig, die die Bereiche Fintech, Gesundheitswesen und Fertigkeiten voranbringen. Derzeit ist er Präsident und Leiter von ChainAim, das technische Strategieberatung sowie Anwendungs- und Entwicklungsdienste anbietet.

Sathya Krishnasamy half bei der Konzeption der Verwendung von ICHOM für Daten, bei der Bewertung des synthetischen Datensatzes, bei der Erstellung einer Basislinie für die Modellkomplexität und bei der Bewertung des ZKML-Anwendungsfalls im Gesundheitswesen.

Dr. Govindarajan ist der Chief Medical Officer von GuardianMedX. Er ist eine Führungskraft im Gesundheitswesen mit einem starken medizinischen Hintergrund und technologischen Kenntnissen. Er verfügt über 35 Jahre umfassende Erfahrung in der internen

Medizin und Geriatrie in Indien und ist als Berater für geriatrische und palliative Versorgung für viele staatliche Einrichtungen in Indien tätig. Er hat in Kliniken, Krankenhäusern, Pflegeheimen, Hospizen und Heimen eine patientenzentrierte Qualitätspflege nach einem einzigartigen Pflegekonzept geleitet und verwaltet.

Dr. Govindarajan hat die Initiative ins Leben gerufen und die Forschung für die Datenerfassung, die Entwicklung und Auswertung des ICHOM für Diabetesdaten sowie die klinische Auswertung der synthetischen Datensätze durchgeführt.

Datenverfügbarkeitserklärung (DAS), gemeinsame Nutzung von Daten, Reproduzierbarkeit und Datenrepositories

Das Datenwörterbuch für das ICHOM V5 Diabetes-Datenwörterbuch ist unter https://www.ichom.org/ patient-centered-outcome-measure/diabetes/ verfügbar.

Anwendung von KI-generiertem Text oder verwandter Technologie

Keine.

Danksagungen

Yugesh Panta, ein Master of Science Student am Depart-ment of Electrical and Computer Engineering, Tandon School of Engineering, New York University, half den Auftraggebern bei der Sammlung von Forschungsdaten, der Validierung, der Analyse und der Modellabstimmung.

Referenzen

1. Goldwasser S, Micali S, Rackoff C. The knowledge complexity of interactive proof systems. Proceedings of the seventeenth annual ACM symposium on Theory of computing-STOC '85. 1985.
2. Fiat A, Shamir A. Wie man sich selbst beweist: praktische Lösungen für Identifikations- und Signaturprobleme. Advances in Cryp-tology-CRYPTO' 86 [Internet]. 2019;186-94. Verfügbar unter: https://link.springer.com/chapter/10.1007%2F3-540-47721-7_12 [zitiert 2024 Juli 31].
3. Ben-Sasson E, Chiesa A, Tromer E, Virza M. Succinct non-interactive zero knowledge for a von neumann archi-tecture [Internet]. 2019. Verfügbar unter: https://eprint.iacr. org/2013/879.pdf [zitiert am 31. Juli 2024]
4. Ben-Sasson E, Bentov I, Horesh Y, Riabzev M. Scalable, transpar-ent, and post-quantum secure computational integrity [Internet]. ePrint IACR. 2018. Verfügbar unter: https://eprint.iacr.org/2018/046 [zitiert 2024 Juli 31]
5. Guerra-Manzanares A, Lechuga J, Maniatakos M, Shamout FE. Datenschutzgerechtes maschinelles Lernen im Gesundheitswesen: offene Herausforderungen und Zukunftsperspektiven. ICLR 2023 Workshop über vertrauenswürdiges maschinelles Lernen für das Gesundheitswesen. arXiv. 2023; 1-13. https://arxiv.org/abs/2303.15563
6. Gotor AM. Maximierung der Modelleffizienz mit model-com-plexity-reducer (MCR). zkcook/docs/mcr.pdf at main giza-techxyz/zkcook [Internet]. GitHub. [cited 2024 Aug 1]. Verfügbar unter: https://github.com/gizatechxyz/zkcook/blob/ main/docs/mcr.pdf
7. Das AK, Saboo B, Maheshwari A, Nair VM, Banerjee S, Jay-akumar C, et al. Health care delivery model in India with rele-vance to diabetes care. Heliyon. 2022 Oct;8(10):e10904. https:// doi.org/10.1016/j.heliyon.2022.e10904
8. Musacchio N, Giancaterini A, Guaita G, Ozzello A, Pellegrini MA, Ponzani P, et al. Artificial intelligence and big data in diabetes care: a position statement of the Italian Associa-tion of Medical Diabetologists. J Med Internet Res. 2020 Jun 22;22(6):e16922. https://doi.org/10.2196/16922
9. Diabetes [Internet]. ICHOM. [zitiert 2024 Aug 1]. Verfügbar unter: https://www.ichom.org/patient-centered-outcome-measure/ diabetes/
10. Benning L, Das-Gupta Z, Fialho LS, Wissig S, Tapela N, Gaunt S. Balancing adaptability and standardisation: insights from 27 routinely implemented ICHOM standard sets. BMC Health Serv Res. 2022 Nov 28;22(1):1424. https://doi.org/10.1186/ s12913-022-08694-9

APPENDIX

Definierte Akronyme

AI/ML: Künstliche Intelligenz / Maschinelles Lernen

HBA1c: Glycoxyliertes Hämoglobin

ICHOM: Internationales Konsortium für die Messung von Gesundheitsergebnissen

LightGBM: Light Gradient Boosting Machine MCR: Modellkomplexitätsreduzierer

ML: Maschinelles Lernen

zk-SNARKs: Zero-Knowledge Succinct Non-Interactive Arguments of Knowledge

ZKML: Null-Wissen Maschinelles Lernen ZKP: Null-Wissensbeweise