

Automatisierte Auswertung von Blutproben mittels Machine Learning

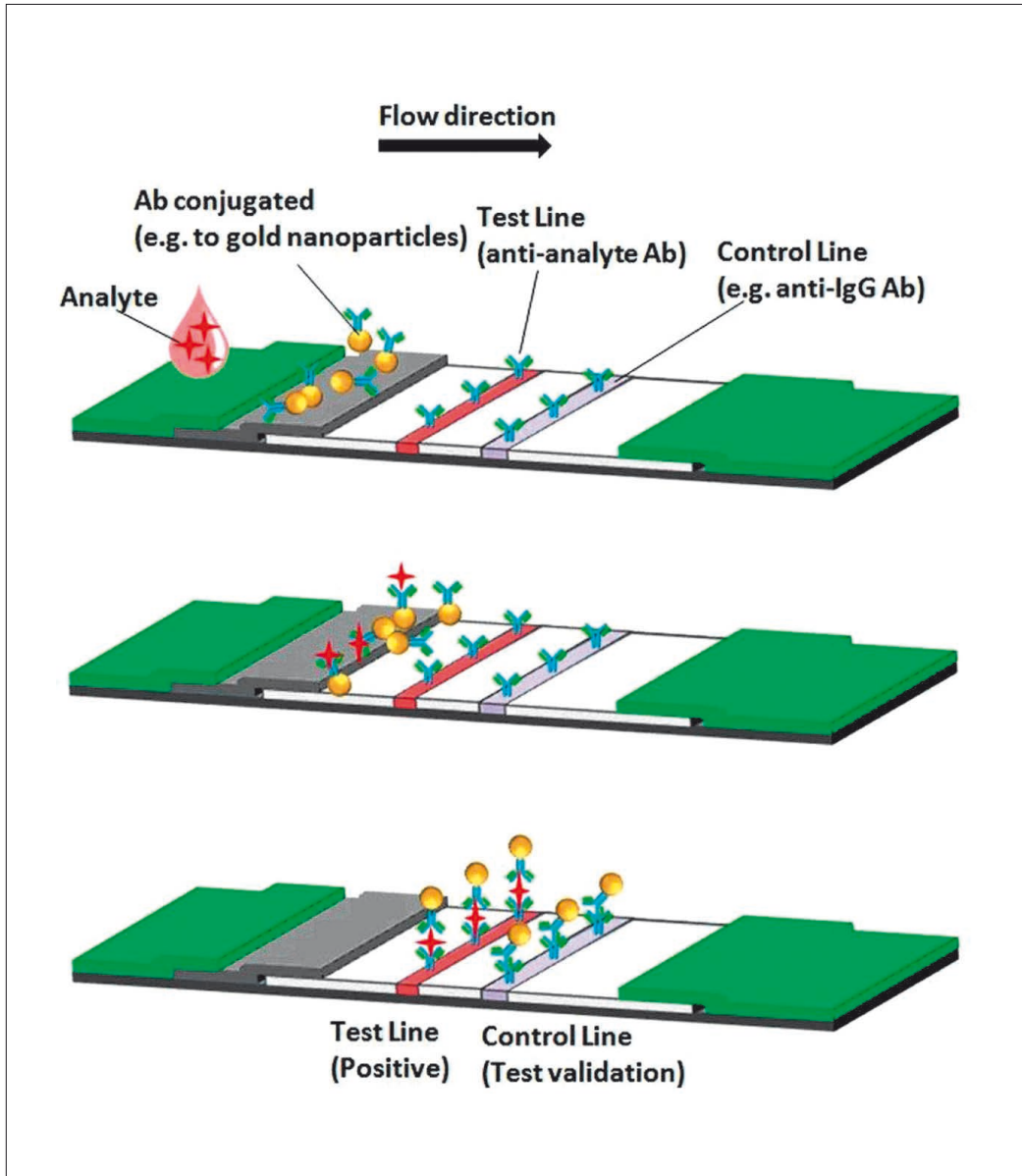


Bild 1: Ablauf eines Lateral Flow Assay [1]

Künstliche Intelligenz hat im Verlauf der Jahre in vielen Bereichen stark an Bedeutung gewonnen. Auch im Medizinbereich nutzen immer mehr Anwendungen künstliche Intelligenz bzw. Machine Learning für präzisere Diagnosen. Gleichzeitig steigt der Bedarf an schnellen und zuverlässigen Analysemöglichkeiten von Blutproben, die patientennah durchgeführt werden können, weltweit stark an. Lateral Flow Assays (LFA) bieten eine ideale Möglichkeit zur schnellen und

einfachen Anwendung im Point-of-Care-Bereich (POC).

Mithilfe von optischen Messsystemen können LFAs quantitativ ausgewertet werden. Das Erkennen von Störungen auf LFAs ist mit herkömmlicher Bildverarbeitung sehr komplex. Cicor hat mithilfe von Bildverarbeitung optische Messungen von LFAs analysiert und verschiedene Machine Learning Modelle trainiert. Durch den Einsatz von Machine Learning wird gezeigt, dass Störungen auf dem

LFA erkannt und entsprechend klassifiziert werden können.

Lateral Flow Assay (LFA)

LFA ist eine papierbasierte Plattform für die Detektion und Quantifizierung von Stoffen in komplexen Mischungen. Es ist eine biochemische Methode zum qualitativen und quantitativen Nachweis spezifischer Antigene und Antikörper. Eine zu untersuchende Probe wird auf einen Teststreifen aus Polymer aufgetragen und fließt dank der Kapillarwirkung durch verschiedene Zonen. Auf diesen Zonen sind Moleküle befestigt, die mit Analyten der Probe wechselwirken können. Die Probe wandert weiter durch den Bindungsfreigabebereich, der Antikörper enthält. Diese sind spezifisch für den Analyten und sind mit gefärbten oder fluoreszierenden Partikeln gebunden. Meist kommen dafür Gold- oder Latexmikropartikel zum Einsatz. Die Probe fließt mit den gebundenen Antikörper entlang des Teststreifens in die Nachweiszone. Dabei handelt es sich um eine poröse Membran mit spezifischen biologischen Komponenten, die in Linien befestigt sind. Eine Reaktion der Antikörper gebundenen Analyten mit dieser Membran führt zu einer Verfärbung der Testlinie. Eine Kontrolllinie dient dazu den korrekten Flüssigkeitsfluss durch den Streifen anzuzeigen (Bild 1).

Machine Learning

Machine Learning (Maschinelles Lernen) kann als Teilgebiet der künstlichen Intelligenz bezeichnet werden, dessen Ziel es ist, bestimmte Muster oder Regeln in einer vorgegebenen Datenmenge automatisiert zu finden. Um dies zu erreichen, werden Lernalgorithmen mit Beispielen trainiert und daraus wird ein komplexes Modell entwickelt. Das Modell kann anschließend auf neue, unbekannte Daten ähnlicher Struktur angewendet werden und somit diese Daten beurteilen.

Es gibt verschiedene Arten von Machine Learning Methoden, die jeweils für unterschiedliche Zwecke

Autor:
Claudio Furter
Software Engineer
Cicor Group
www.cicor.com

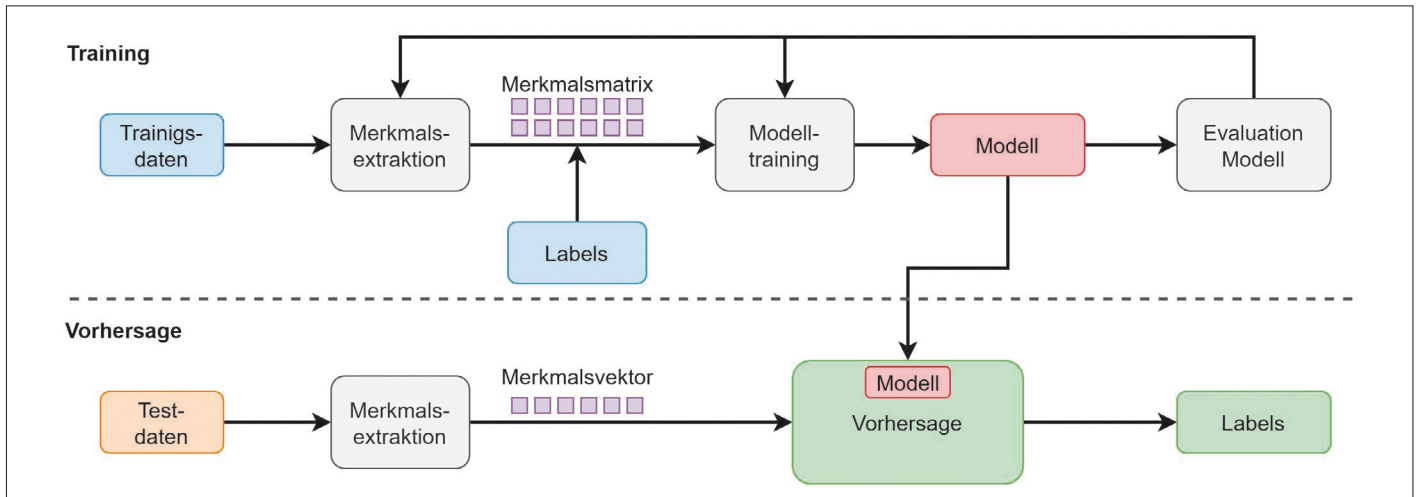


Bild 2: Ablauf von Supervised Learning

geeignet sind. Beim Supervised Learning (Überwachtes Lernen) müssen die richtigen Antworten, genannt Labels, in den Trainingsdaten mitbegriffen sein. Das Unsupervised Learning (Unüberwachtes Lernen) hingegen benötigt lediglich die rohen Trainingsdaten, um Muster in den Daten zu erkennen und dadurch eine Gruppierung der Daten zu erzeugen.

Supervised Learning

Die Herausforderung beim Supervised Learning besteht also darin die Beziehung zwischen den Trainingsbeispielen und Labels zu modellieren. Das daraus resultierende Modell wird benutzt, um für neue und unbekannte Daten ein Label vorherzusagen. Bild 2 stellt die grundlegenden Schritte von Supervised Learning dar. Zuerst werden die entspre-

chenden zuvor definierten Merkmale aus den Trainingsdaten extrahiert und daraus eine Merkmalsmatrix erstellt. Anhand der extrahierten Merkmale und der dazugehörigen Labels wird ein Algorithmus trainiert, sodass ein auf die Trainingsdaten angepasstes Vorhersage-Modell entsteht. Das Modell wird anschließend mit noch unbekanntem Validierungsdaten evaluiert. Je nach Genauigkeit der Vorhersagen muss allenfalls die Merkmalsextraktion neu definiert oder das Modell optimiert werden. Bei Vorhersagen von den ungelabelten Testdaten werden die beim Training definierten Merkmale extrahiert und vom Modell verarbeitet.

Unsupervised Learning

Beim Unsupervised Learning werden Trainingsdaten ohne dazu-

gehörige Labels verwendet. Diese Methode wird meistens eingesetzt, wenn große, unstrukturierte Daten vorhanden sind, von denen man nicht weiß nach welchen Kriterien diese aufgeteilt werden können. Unsupervised Learning Methoden versuchen Strukturen und Unterschiede in den Daten zu erkennen, um Gruppen (Cluster) ähnlicher Beispiele zu finden. Es wird ebenfalls ein Modell gebildet, welches zur Vorhersage der Gruppen dient.

Für Machine Learning gibt es eine Vielzahl von Modelltypen und Algorithmen, welche jeweils für unterschiedliche Aufgaben geeignet sind. Die Aufgabe, störbehaftete LFAs zu erkennen, ist ein Klassifikationsproblem. Es werden deshalb verschiedene Klassifikationsverfahren des Supervised Learning für die Problemlösung eingesetzt.

die Trainingsdaten erneut in Trainingsdaten und Validierungsdaten aufgeteilt. Mittels Kreuzvalidierung können so Machine-Learning-Algorithmen trainiert und validiert werden. Das entstandene Modell wird anschließend auf die bisher noch unbekanntem Testdaten angewendet und somit kann eine Aussage über die Performance des Modells getroffen werden.

Vorbereitung zum Trainieren

Vor dem Trainieren der Algorithmen müssen die Graustufen-Bilder in numerische Eingangsdaten umgewandelt werden. Dafür werden die Bilder mit Methoden der Bildverarbeitung entsprechend aufbereitet und unterschiedliche statistische Merkmale extrahiert. Diese extrahierten Merkmale werden in einer Matrix abgespeichert und stellen die Grundlage für das Trainieren der Machine-Learning-Modelle dar. In einem nächsten Schritt werden verschiedene Machine-Learning-Algorithmen wie beispielsweise Klassifikationsbäume oder Support Vector Machine miteinander verglichen und deren Vorhersage-Performance betrachtet. Dieser erste Vergleich zeigt, dass mit den extrahierten Merkmalen und den Algorithmen k-Nearest-Neighbor und Random Forest bereits ein akzeptables Resultat erreicht werden kann.

Training und Evaluation von Machine-Learning-Modellen

Als Datengrundlage für das Trainieren der Machine Learning Modelle wird ein Datensatz von optischen LFA-Messungen in Form von Graustufen-Bildern verwendet. Die Daten sind in drei Klassen bzw. Labels eingeteilt, wobei eine Klasse die störungsfreien LFAs beinhaltet sowie zwei unterschiedliche Fehlerklassen mit störungsbehafteten LFAs (Bild 3).

Der gesamte Datensatz wurde für das Trainieren und Validieren der Algorithmen zufällig in 80 % Trainingsdaten und 20 % Testdaten aufgeteilt. Die drei Klassen sind verhältnismäßig gleich verteilt vertreten. Gemäß Bild 4 sind

Parameter-Tuning

Um die Performance der Modelle weiter zu steigern, wird ein Parameter-Tuning durchgeführt. Die Machine-Learning-Algorithmen haben verschiedene einstellbare

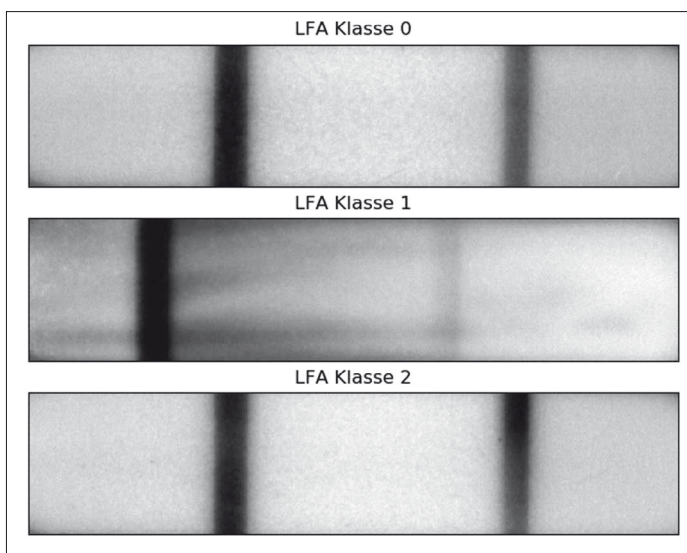


Bild 3: Störungsfrei: Klasse 0, Störungsbehaftet: Klasse 1 und 2

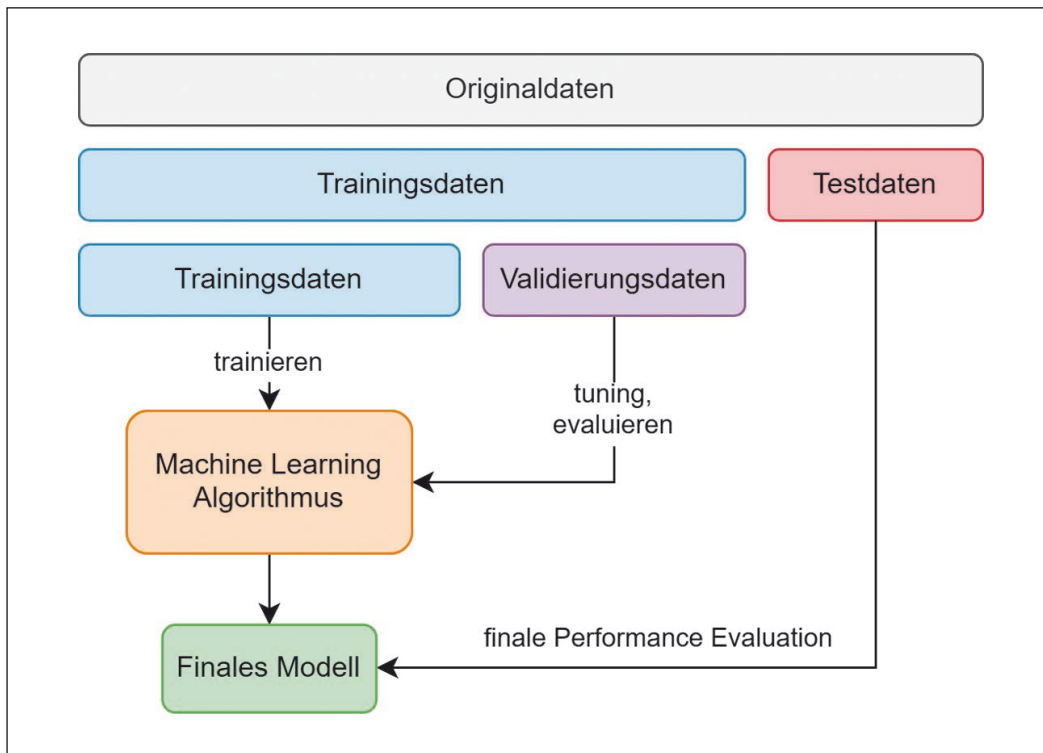


Bild 4: Ablauf Training, Evaluation und Validierung

Parameter. Mit Parameter-Tuning kann die optimale Einstellung dieser Parameter automatisch ermittelt werden. Dabei wird pro Parameter Kombination ein Modell trainiert und validiert. Das Modell mit der besten Performance hat somit die optimalsten Parameter.

Oversampling-Methoden

Aufgrund einer großen Klassenungleichheit in den Daten werden zusätzlich verschiedene Oversampling-Methoden zur Verbesserung der Performance eingesetzt. Mit Methoden wie Random Oversampling oder SMOTE (Synthetic Minority Oversampling Technique) werden die Minderheitsklassen auf verschiedene Arten vergrößert und sollen so stärker gewichtet werden. Diese

beiden Methoden werden vorgängig zum Parameter-Tuning angewendet.

Es hat sich gezeigt, dass der Random Forest unter Anwendung von SMOTE in dieser Anwendung die beste Performance erreicht. Die Qualität des Modells wird mit dem F1-Score gemessen. Der F1-Score ignoriert die Klassenungleichgewichte und ist für diesen Fall gut geeignet.

Das finale Modell erreicht einen F1-Score von 82 % auf den Testdaten. Basierend auf dem kleinen Datensatz ist dies bereits ein gutes Ergebnis. Die Auswertung zeigt, dass vor allem LFAs mit extremen Störungsfällen zuverlässig und richtig klassifiziert werden. Mit zusätzlichen Daten könnte die Performance des Modells gesteigert

und somit Grenzfälle besser trainiert werden.

Machine Learning im Medizinbereich

Damit das evaluierte Modell auf einem Medizingerät eingesetzt werden kann, gilt es einige regulatorische Anforderungen und allgemeine Normen zu beachten. In der traditionellen Entwicklung einer Software bzw. eines Programmes wird eine gewünschte Berechnung oder ein Problem von einer Person gelöst und ein entsprechender Code implementiert. Dieses Vorgehen zur Problemlösung und Entwicklung eines Programmes verläuft nach einem definierten Prozess. Ist das Programm für den Betrieb freigegeben, kann es auf der Ziel-

plattform eingesetzt werden und eine bestimmte Eingabe führt zu einem erwarteten Ergebnis. Durch den Einsatz von Machine Learning verändert sich dieser Ablauf. Eine Berechnung oder ein Problem wird neu durch einen Machine-Learning-Algorithmus gelöst. Das daraus resultierende Modell wird dann auf der Zielplattform eingesetzt und kann Vorhersagen treffen.

Traditionelle vs. Machine Learning Entwicklung

Bild 5 vergleicht die traditionelle Entwicklung mit der Machine Learning Entwicklung. Aktuell existieren keine Gesetze und harmonisierte Normen, welche den Einsatz von künstlicher Intelligenz bzw. Machine Learning in Medizinprodukten regulieren. Es ist deshalb wichtig die allgemeinen Normen und Best Practices einzuhalten. Trotz der nicht existierenden Normen für den Einsatz von Machine Learning in Medizinprodukten müssen die bestehenden regulatorischen Anforderungen wie die MDR und IVDR erfüllt werden. [2]

Zusammenfassung

Bereits heute wird künstliche Intelligenz in Medizinprodukten für Diagnosen eingesetzt und wird in Zukunft weiter an Bedeutung gewinnen. Durch die Integration des erstellten Modells auf einem optischen Messsystem können Messergebnisse von Blutproben verbessert werden indem stöberhafte Messungen zuverlässiger erkannt werden können.

Quellen

- [1] <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4986465/>
- [2] Johner Institut, Seminar „KI-Anwendungen gesetzeskonform entwickeln“ ◀

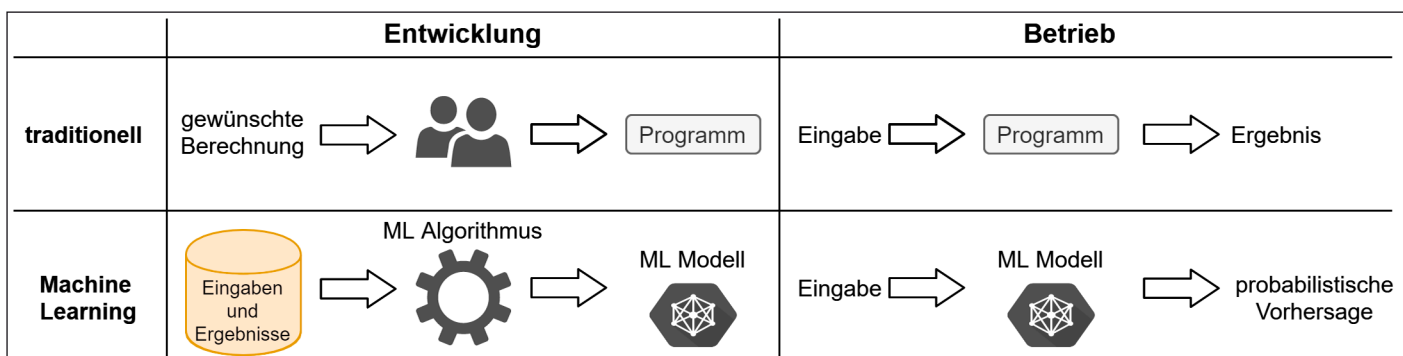


Bild 5: Traditionelle vs. Machine Learning Entwicklung