

Einführung in die Forschungsmethoden der Psychologie und Psychotherapie

Einheit 9: Digitale Datenerhebung, Big Data und Machine-Learning

04.07.2024 | Dr. Caroline Zygar-Hoffmann

Heutige Themen

Digitale Datenerhebung

Ambulatory Assessment & Ecological Momentary Interventions

Big Data und Machine-Learning

Take-Aways und Schlüssel-/Fachbegriffe

- Take-Aways
- Schlüssel-/Fachbegriffe

Digitalisierung traditioneller Datenerhebungsmethoden

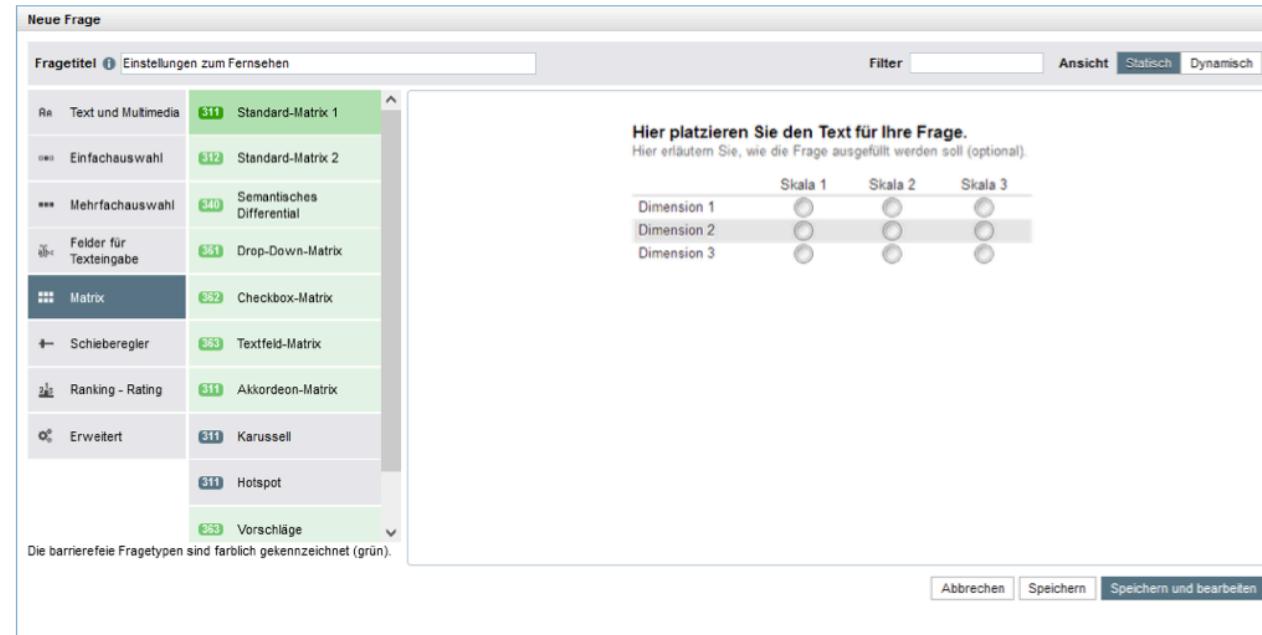
- Fragebögen und Tests wurden früher im Paper Pencil Format vorgegeben
- Mittlerweile ist die digitale Form vorherrschend

Methode: Surveytools

- Digitalisierung der Messinstrumente durch die Forscher:innen
 - Forschungsmethodische Aspekte (z.B. Randomisierungslogik) wird miteinbezogen (programmiert)
 - Ausfüllen durch Proband:innen am Computer/Smartphone
 - Nach Ende der Datenerhebung → direkter Export der Daten ins digitale Format
- Erstellen digitaler Umfragen wird im Modul Empirisch-wissenschaftliches Arbeiten gelernt (2. Semester)

Digitalisierung traditioneller Datenerhebungsmethoden

Surveytools - Beispiel Unipark:



→ recht benutzerfreundlich, aber kostenpflichtig

Digitale Datenerhebung

Digitalisierung traditioneller Datenerhebungsmethoden

Surveytools - Beispiel formR:

Notes	class	type	option_name	showif	label	choice1	choice2
	mc_width80 space_bottom_5	mc_heading mc_aut		AntwortSkala_AutonomieSkala			
	mc_width80	rating_button 1,5,1	item1_AutonomieSkala		In unserer Partnerschaft kann ich ohne größere Einschränkungen machen, was ich will.		
	mc_width80	rating_button 1,5,1 submit	item2_AutonomieSkala submit4		In unserer Partnerschaft kann ich meinen eigenen Interessen nachgehen, ohne dass mein Partner sauer ist. Weiter		
Rückmeldungsfrage	left500	note mc	rueckmeldung mc_rueckmeldung	### <center> Rückmeldung </center> Möchten Sie eine Rückmeldung zu ihren eigenen Daten erhalten?		ja	nein

Diagramm zur Konfiguration der Rückmeldungsfunktion:

list_name	name	label
mc_aut	1	Trifft überhaupt nicht zu
	2	
	3	
	4	
	5	Trifft voll und ganz zu

→ Surveys werden über Exceltabelle konfiguriert, mächtig und dadurch sehr flexibel, versteht R Befehle, kostenfrei

Möglichkeiten und Vorteile digitaler Datenerhebung

Möglichkeit Daten unabhängig von der physischen Distanz zwischen den Forschungseinrichtungen und den Teilnehmer:innen zu erheben

Beispiele für Vorteile sind:

- Zeitökonomie
- Keine Übertragungsfehler von Papier in den Computer
- Direkter Export der Antworten ins Datensatzformat, d.h. Analyse kann direkt starten
- Wegfall von kopräsenten Versuchsleitung und Versuchsleitereffekten
- Verringerung des Reaktivitätsproblems
- größere Stichprobengrößen erreichbar und damit höhere Teststärke
- Diversifizierung der Stichproben indem internetbasiert rekrutiert wird
- Erreichbarkeit von Stichproben mit hochspezifischen Merkmalen

Nachteile digitaler Datenerhebung

Digitale Datenerhebung bietet der psychologischen Forschung nicht nur Möglichkeiten und Vorzüge, sondern ist auch mit Risiken und potenziellen Nachteilen behaftet.

Beispiele für Nachteile sind:

- Verringerung der Repräsentativität der Stichproben hinsichtlich Technologieaffinität
- Erschwerte Kontrolle über die Bedingungen der Datenerhebung (z.B. die Einhaltung von Instruktionen, Einhalten der Datenqualität)
- Erschwerter Schutz der Teilnehmenden
- Erschwerte Überprüfung des Verständnisses wichtiger Informationen
- Erschwerte Prüfung der Identifizierbarkeit
- erschwerter Datenschutz und erschwerte informierte Einwilligung

Ambulatory Assessment & Ecological Momentary Interventions

Was ist Ambulatory Assessment?

- **Wiederholte Erfassung** der aktuellen Gedanken, Emotionen, Verhaltensweisen, physiologischen Zustände und des Kontexts von Menschen **in ihrer natürlichen Umgebung** und möglichst im Moment in dem Sie auftreten ("Real-time Data Capturing" bzw. "Real-time Monitoring")
- Typischerweise (aber nicht notwendigerweise) über elektronische tragbare Geräte
- Der Begriff "Ambulatory Assessment" (Ebner-Priemer & Trull, 2009) umfasst eine Reihe methodischer Ansätze
 - "Experience Sampling Method" (ESM, Larson & Csikszentmihalyi, 1983), auch "Ecological Momentary Assessment" (EMA, Shiffmann et al., 2008)
 - "Daily Diaries" (Gunther & Wenzel, 2012)
 - "Mobile Sensing"
 - "Ambient Sensing"
- Die verschiedenen Begriffe betonen unterschiedliche Aspekte der Forschung (z.B. was wird erfasst, wie und wie oft wird erfasst) und werden auch in verschiedenen Subdisziplinen häufiger oder weniger häufig genutzt

ESM/EMA Sampling Pläne

ESM/EMA kann zwei unterschiedlichen Messregimen folgen:

1) Time-contingent sampling bzw. Signal-contingent sampling

Teilnehmer antworten auf Fragen zu einem bestimmten Zeitpunkt, wenn sie durch ein Signal dazu aufgefordert werden ("Beeps")

- **Fixed scheme:** Teilnehmer werden zu festen Zeitpunkten aufgefordert, z. B. um 9:30 und 16:30 Uhr → für die Versuchspersonen vorhersehbar, höhere Antwortquote, aber höhere Reaktivität
- **Random scheme:** Aufforderungen werden zu zufälligen Zeitpunkten gesendet, aber i.d.R. in vorher festgelegten Intervallen → bessere Abdeckung des gesamten Alltags, ggf. geringere Reaktivität, aber auch geringere Antwortquote

Ambulatory Assessment & Ecological Momentary Interventions

ESM/EMA Sampling Pläne

ESM/EMA kann zwei unterschiedlichen Messregimen folgen:

2) Event-contingent sampling

Teilnehmer antworten auf Fragen, wenn ein bestimmtes Ereignis eintritt, z.B. eine Panikattacke oder Alkoholkonsum

- Möglichkeit 1: Teilnehmer bitten, Fragebogen selbst zu starten sobald das Ereignis eintritt
- Möglichkeit 2: Kombination ESM/EMA mit Sensing:
 - z.B. Stellen der Frage, wenn bestimmtes Aktivitätslevel überstiegen wird
 - z.B. Stellen der Frage, wenn bestimmte GPS Location erreicht wird

Ambulatory Assessment & Ecological Momentary Interventions

ESM/EMA: Beispiele

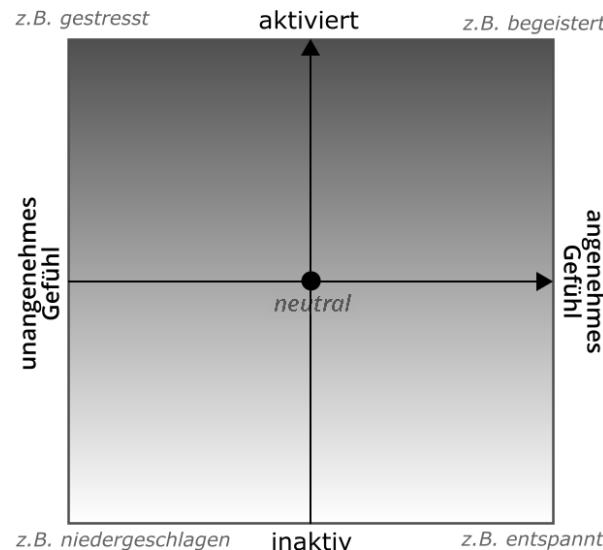
The figure displays three mobile application screens for Ecological Momentary Assessment (EMA) or Experience Sampling Method (ESM). Each screen has a blue header with a back arrow and an information icon.

- Wunsch: Interessen (Left Screen):** A text-based question asking if the user wishes to pursue their own interests. Below it is a Likert scale from "ja, sehr stark" (yes, very strong) to "nein, das würde mich gerade stören" (no, that would bother me just now).
- To-Do-Liste (Middle Screen):** A text-based question asking about the amount of work awaiting the user in the next two hours. Below it is a Likert scale from "gar keine" (none at all) to "sehr viel" (very much).
- Situationen (Right Screen):** A text-based question asking if certain situations have occurred since the morning. Below it is a list of eight checkboxes for various relationship situations.

Ambulatory Assessment & Ecological Momentary Interventions

ESM/EMA: Beispiele

Multi-dimensionale Stimmungsmessung



Zirkumplexmodell des Affekts (Russell, 1980; Russell et al. 1989)

- Alle Stimmungszustände sind lineare Kombination aus zwei unabhängigen, bipolaren Skalen:
 - Valenz (von unangenehm bis angenehm)
 - Erregung/Aktivierung (von niedriger bis hoher Erregung).
- Kombination ordnet die affektiven Zustände auf einem von vier Quadranten an (\rightarrow) Messung über ein "Affect Grid" (Bild links)
- Nachteil: Konzepte von Valenz und Erregung können den Befragten schwer zu vermitteln sein

Ambulatory Assessment & Ecological Momentary Interventions

Was sind Ecological Momentary Interventions (EMI)?

- EMI: **mobilbasierte Interventionen**, die den Menschen in ihrem **Alltag**, also in ihrer natürlichen Umgebung, und auf einen **konkreten Moment/Tag** bezogen angeboten werden (in "Echtzeit")
- Unterschied zu IMI (vgl. Vorlesung von Patricia Garatva) liegt vor allem im *momentary*-Teil, d.h. dass die Interventionen an bestimmten Momenten/Tagen vorgegeben werden
- Begriff auch eher in der Forschung üblich als in der psychotherapeutischen Praxis
- Personalisierung möglich hinsichtlich Timing und Art der Intervention:
 - **Timing:** z.B. wenn ESM/EMA deutliche Veränderungen in relevanten pathologischen Prozessen widerspiegelt
 - **Art:** Auf eine Art und Weise, die in Anbetracht dessen, was über die Person bekannt ist, maximal effektiv ist

Ambulatory Assessment & Ecological Momentary Interventions

Beispiel für Einsatz von EMI in der Partnerschaftsforschung

Eliciting Short-Term Closeness in Couple Relationships With Ecological Momentary Interventions

Caroline Zygar-Hoffmann¹ ,^a Lara Cristoforo¹, Lisa Wolf¹, Felix D. Schönbrot¹ 

¹ Department of Psychology, Ludwig-Maximilians-Universität München, Munich, Germany

Keywords: state relationship closeness, intimacy, ecological momentary interventions, couples, experience sampling method

<https://doi.org/10.1525/collabra.38599>

Collabra: Psychology

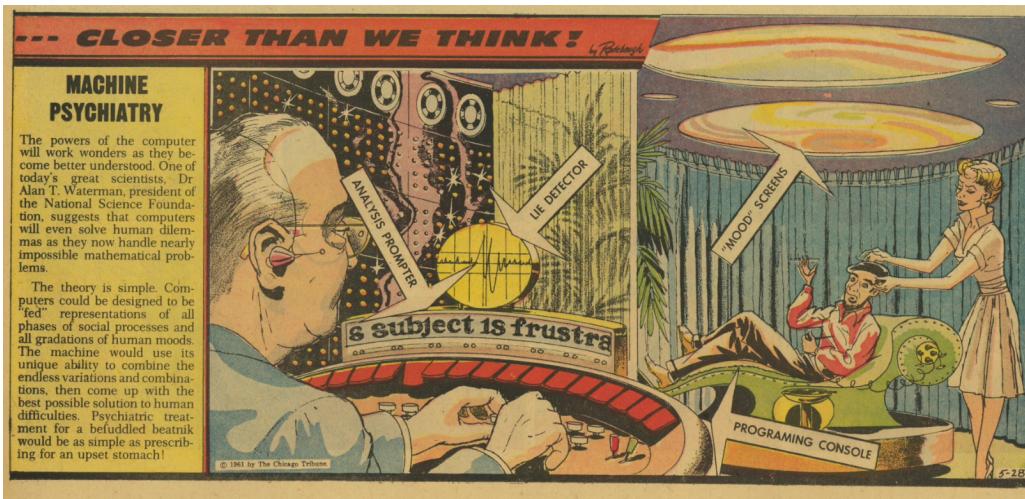
Vol. 8, Issue 1, 2022

Ein Beispiel-EMI aus dem Paper:

"Nehmen Sie sich heute in einem ruhigen Moment etwas Zeit zusammen mit Ihrem Partner. Stellen Sie einen Wecker auf 5 Minuten und schauen Sie Ihrem Partner für 5 Minuten in die Augen. Alternativ können Sie auch Videochatten. Konzentrieren Sie sich dabei voll und ganz auf Ihren Partner."

Computer und mentale Gesundheit

- Idee, Computer einzusetzen, um Diagnostik und Prognostik von Erkrankungen zu verbessern, ist nicht neu



Arthur Radbow (60er Jahre): Machine Psychiatry aus „Closer than we think“.

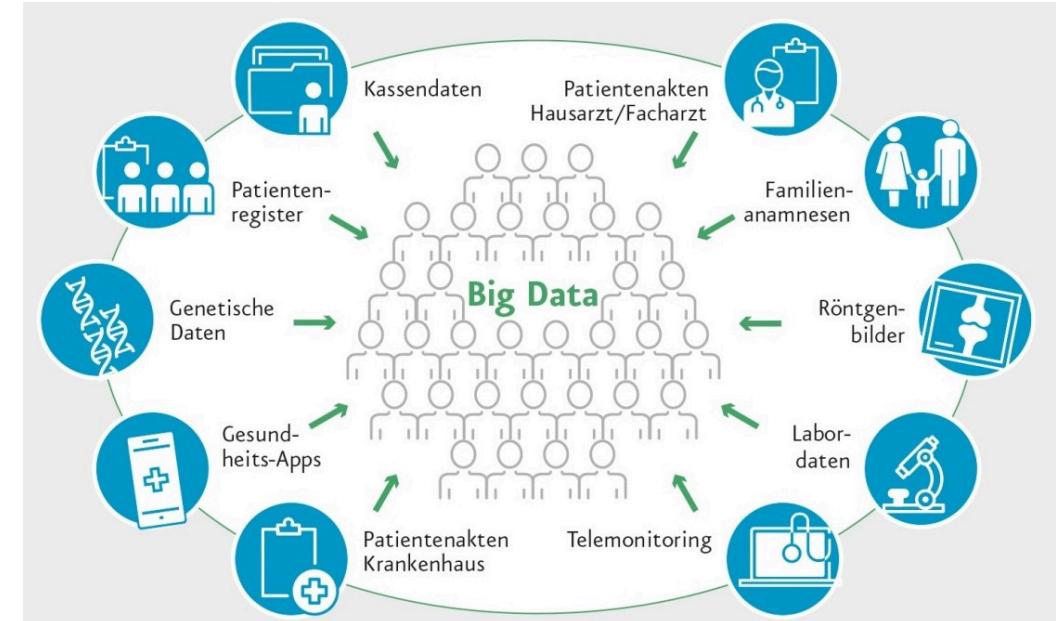
- Idee: einen Computer zu füttern mit allen Wesensäußerungen des Patienten.
- Computer lernt aus den unterschiedlichsten Kombinationen von emotionalen Ausdrücken, Kognitionen und Verhalten, Entscheidungsregeln abzuleiten
- Die Kliniker:in kann diese verwenden, um Therapie danach auszurichten.

- Dichte an Informationen, die wir über Erkrankungen haben – nicht nur über psychische Erkrankungen, sondern generell über Erkrankungen – nimmt exponentiell zu
- **Kapazitätsproblem:** menschliche Kapazität mit diesem Wissen umzugehen ist begrenzt (Schere verfügbares vs. kognitiv verarbeitbares Wissen)

Big Data und Machine-Learning

Big Data

- **Definition:** große Datenmengen, die zu komplex sind, um sie mit traditionellen Methoden zu verarbeiten.
- umfasst sowohl strukturierte als auch unstrukturierte Daten, wie z.B. Text, Bilder, Audio- und Videodateien.
- Neue Technologien ermöglichen es, big data zu speichern und zu verarbeiten.
- **Ziel:** Muster, Trends und Verbindungen in den Daten zu erkennen und zu nutzen.



Machine Learning

- Computerstrategie, die **automatisch** Methoden und Parameter bestimmt (d.h. lernt), um eine optimale Lösung für ein Problem zu finden
- Wird nicht von vornherein von einem Menschen programmiert, um eine fixe Lösung zu liefern.
- Wird als Teilgebiet der künstlichen Intelligenz (KI) betrachtet
- Lernprozess simuliert vermeintlich eine Facette der menschlichen Intelligenz
- Kann für vermeintlich intelligente Zwecke eingesetzt werden (z.B. Sprache, Schrift, Gesichtserkennung, selbstfahrende Autos oder medizinische Entscheidungshilfen).

Machine Learning - die Anfänge

Das Perzeptron:

- Ziel: eine Maschine zu bauen, die ihre eigenen Formeln entwickelt, um Probleme durch Lernen zu lösen (Rosenblatt, 1958)
- Aufgabe: Bilder von Dreiecken von Quadraten unterscheiden mittels 400 Lichtsensoren
- Unterscheidung von Dreiecken von Quadraten nicht expliziert programmiert → Training über **Trial und Error** Prozess



→ In den Anfängen des maschinellen Lernens ging es also darum, dass eine Maschine mit Hilfe realer Daten lernt begrenzte Vorannahmen darüber zu treffen, wie diese Daten aussehen, und dass die Leistung der Maschine unter realen Bedingungen bewertet wird.

<!--

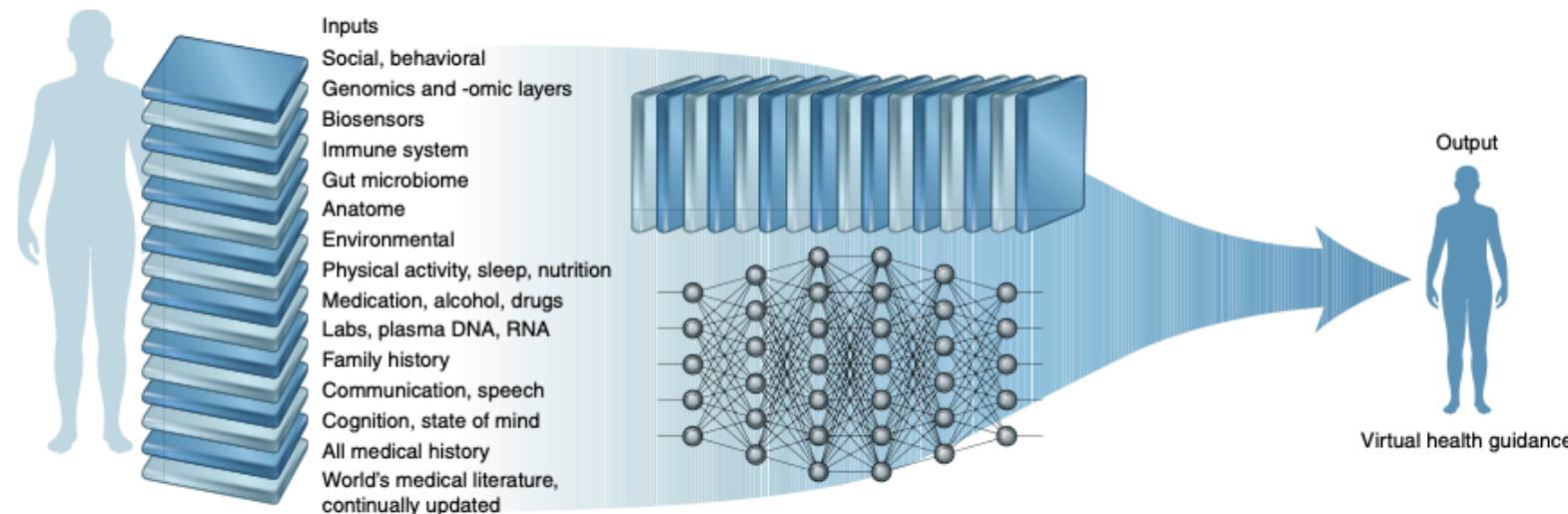
--> <!-- --> <!-- -->

Machine Learning - die Anfänge

- Das Perzeptron trug dazu bei, die Betrachtungsweise von Computern zu ändern
- Vorher: Rechenmaschinen, die von vornherein mit bereits bekannten Formeln und Regeln programmiert werden mussten
- Nachher: Computer, die unscharfe Regeln selbstständig erlernen konnten (z. B. Erkennung von Mustern).
- Dieser frühe Erfolg löste einen großen Hype aus, der größtenteils aus phantasievollen Verallgemeinerungen der zugrundeliegenden Logik auf andere Probleme bestand
- Die Seifenblase platzte jedoch, als sich die Hoffnungen aufgrund der Beschränkungen der frühen Algorithmen und der damals verfügbaren Rechenleistung nicht erfüllten.
- Einige dieser Beschränkungen können nun aufgrund der gestiegenen Rechenleistung überwunden werden, so dass es Maschinen möglich ist aus manchmal sehr komplexen Daten zu lernen.
- Machine Learning Algorithmen sind heute in unser tägliches Leben integriert: Internet-Suchen, Produktempfehlungen, Spracherkennungsdienste, selbstfahrende Autos, Bildklassifikation, Sprachmodelle (z.B. DeepL), Generative Modelle (z.B. ChatGPT)

Machine Learning - Entwicklungen im Gesundheitswesen

- Hoffnung, auch klinische Entscheidungsfindung durch Informationsverarbeitung zu verbessern
- **Präzisionsmedizin (taylored treatments)**: Richtig Behandlung für richtige Patient:in zur richtigen Zeit



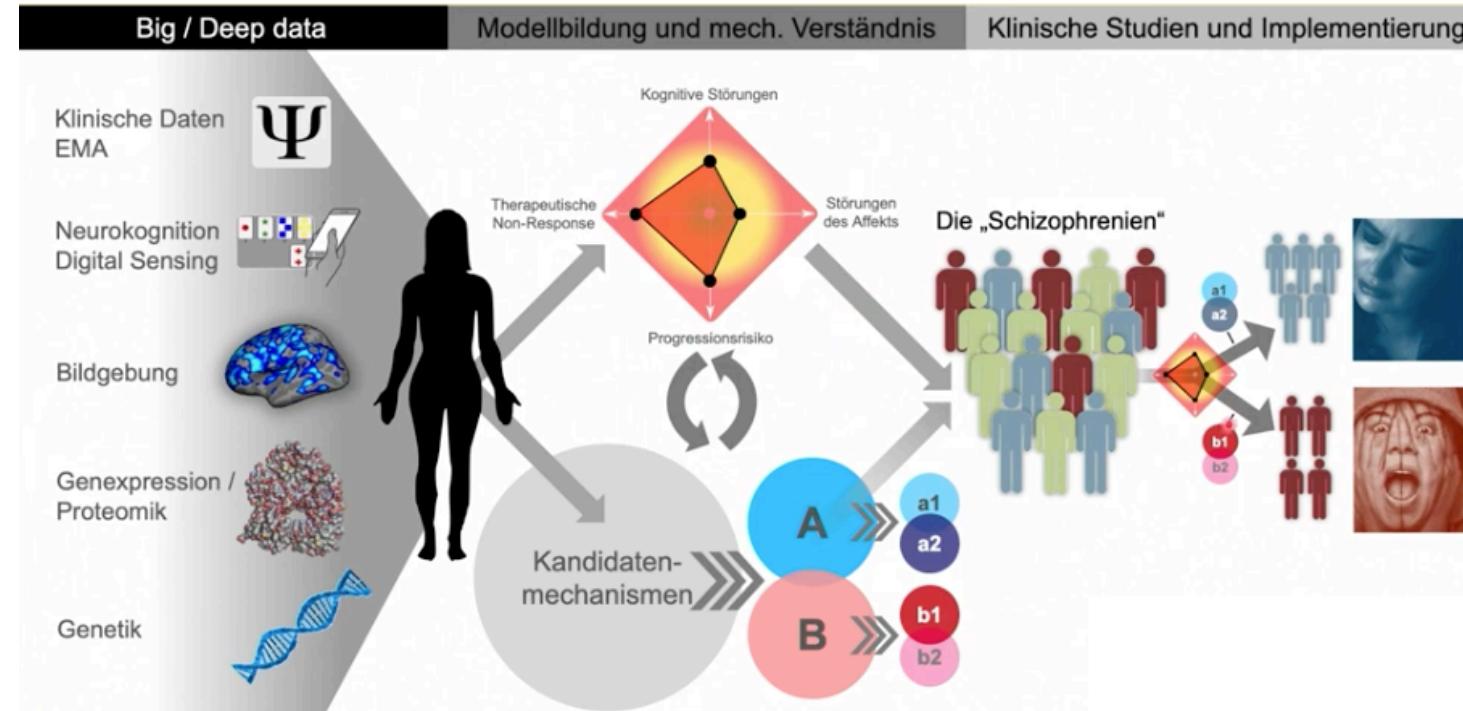
Machine Learning - Entwicklungen im Gesundheitswesen

- Ansatz des maschinellen Lernens im Gesundheitswesen funktioniert insbesondere gut bei Aufgaben, bei denen es um **Mustererkennung in Bildern** geht
- ML ebenso gut oder besser abgeschnitten als Kliniker, etwa bei
 - Erkennung von Hautkrebs (Esteva et al. 2017)
 - Lungenkrebs (Yu et al. 2016)
 - Brustkrebs (z. B. Samsung, RS80A)
 - Augenerkrankungen (Long et al. 2017)
- Und in der Psychologie? **Besonders für das Ziel der Vorhersage sehr relevant!** (für das Ziel der Erklärung weniger) → vielversprechende Forschung

Big Data und Machine-Learning

Machine Learning - Entwicklungen im Gesundheitswesen

Psychologisches Beispiel: Vorhersage von Schizophrenie-Onset im Prodrom (Koutsouleris et al. 2021)



Klassen von Machine Learning

Man unterscheidet im Machine Learning i.d.R. drei unterschiedliche Anwendungsbereiche:

- **Unsupervised Learning:**

- Beinhaltet keine vorgegebenen Labels oder Zielvariablen.
- Ziel ist es, unbekannte Strukturen oder Muster in den Daten zu entdecken.
- Beispiele: Clustering, Dimensionality Reduction.

- **Supervised Learning:**

- Beinhaltet vorgegebene Labels oder Zielvariablen.
- Ziel ist es, ein Modell zu entwickeln, das in der Lage ist, neue Datenpunkte richtig vorherzusagen. → "**Prädiktive Modellierung**"
- Beispiele: Regression, Klassifikation.

- **Reinforcement/Reward Learning:**

- Beinhaltet ein Agent-System, das in einer Umgebung agiert und Entscheidungen trifft.
- Ziel ist es, das Verhalten des Agenten so anzupassen, dass es möglichst hohe Belohnungen erhält.
- Beispiele: Game-playing, Robotics.

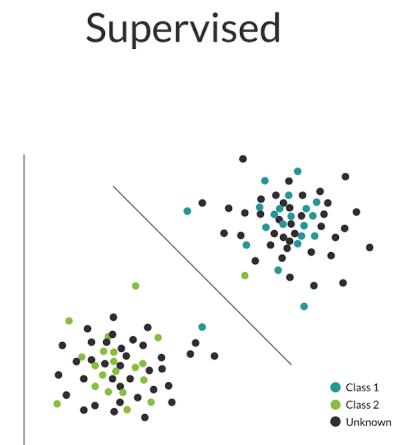
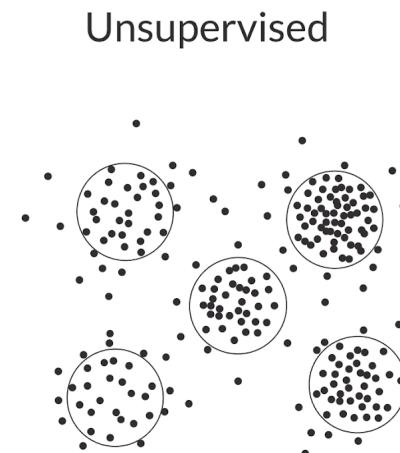
→ Es ist zu beachten, dass diese drei Arten von Machine Learning-Methoden nicht immer klar voneinander abgegrenzt sind und es auch überlappende Anwendungen gibt.

Big Data und Machine-Learning

Klassen von Machine Learning

Unsupervised Learning und Supervised Learning

- Lassen sich noch auf klassische statistische Fragestellungen zurückführen
- Output dieser Algorithmen ist eine Kategorie, bzw. eine Zahl
- **Unsupervised Learning:** Gruppierungsvorschläge für chaotische Datenstrukturen, ohne Vorkenntnis von Kategorien (Clustering) → wichtig für Entdeckung neuer Krankheitsmuster etc.
- **Supervised Learning:** versucht bereits bekannte Kategorien in Daten anhand anderer Informationen in den Daten wiederzuerkennen/vorherzusagen → wichtig für klinische Prognosen/Diagnosen

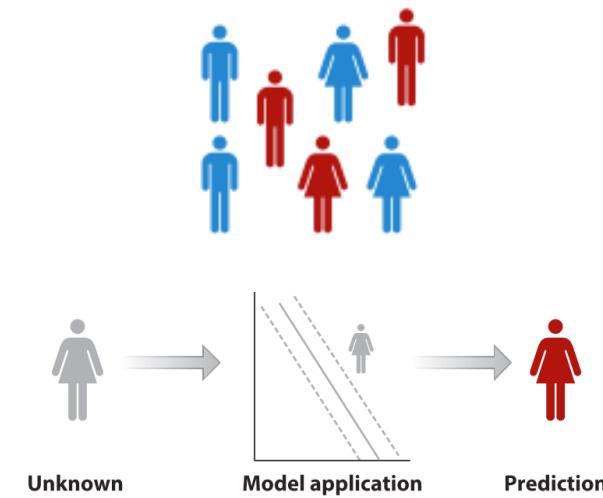


Klassen von Machine Learning

Logik der Prädiktiven Modellierung

Szenario:

- Wir haben einen Trainingsdatensatz mit Brainscans
- Wir wissen für Personen im Trainingsdatensatz, wer eine Diagnose hatte (rot) und wer nicht (blau)
- Prädiktion: Für neue Person (nicht im Datensatz) auf Basis der Brainscans vorhersagen, ob Diagnose vorhanden

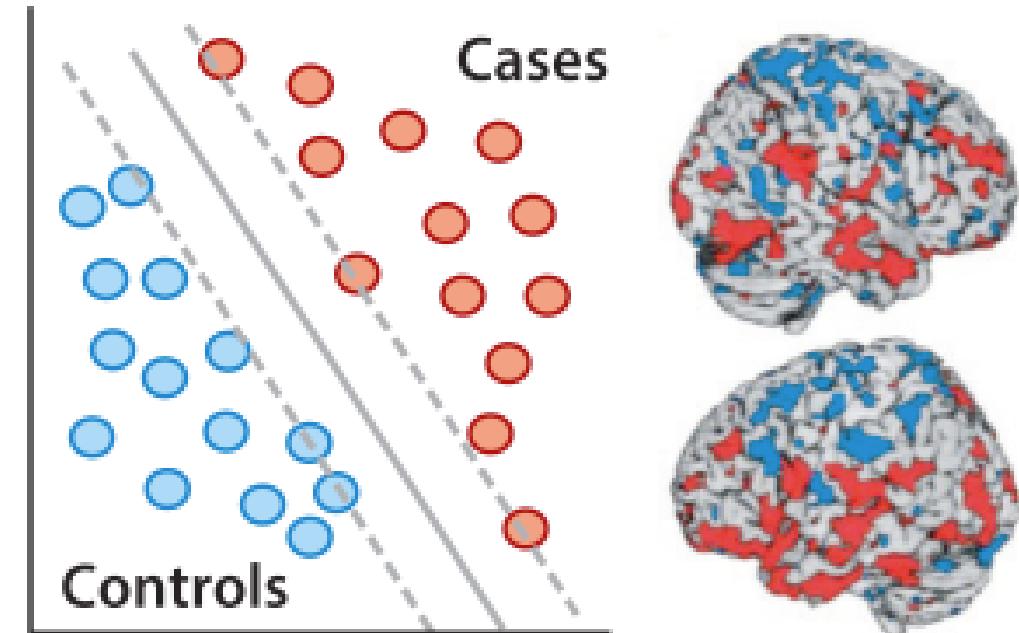


Klassen von Machine Learning

Logik der Prädiktiven Modellierung

Schritt 1:

- Im Trainingsdatensatz das Modell auf Basis der Aktivierungsmuster im Brainscan trainieren
- Das heißt, das Modell versucht anhand der vorhandenen Daten herauszufinden woran es Personen mit und ohne tatsächliche Diagnose unterscheiden kann



Klassen von Machine Learning

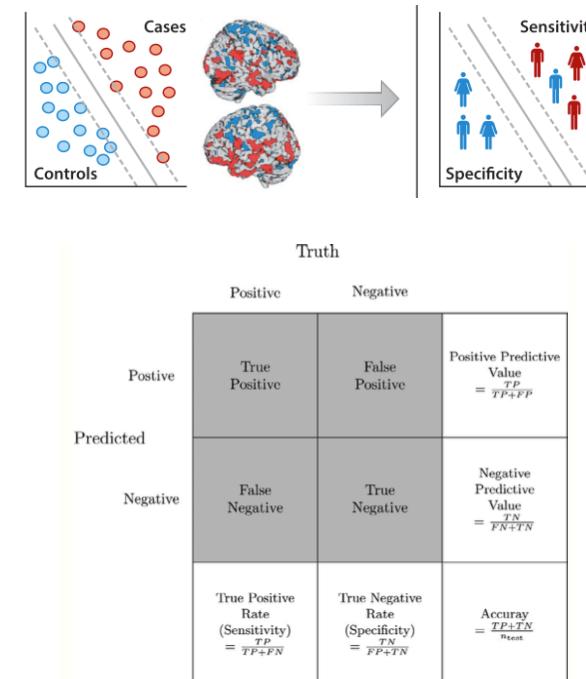
Logik der Prädiktiven Modellierung

Schritt 2:

- Prüfen, ob durch das Modell die tatsächliche Diagnose von Personen **im Trainingsdatensatz** vorhersagen kann (d.h. ob in den Daten überhaupt Informationen vorhanden sind, anhand derer eine Vorhersage möglich ist) .

Genauigkeitsmaße

- gesamte Vorhersagegenauigkeit = **Accuracy**
- Vorliegen von tatsächlicher Diagnose korrekt erkennen = **Sensitivität**
- Abwesenheit von tatsächlicher Diagnose korrekt erkennen = **Spezifität**



Klassen von Machine Learning

Logik der Prädiktiven Modellierung

Schritt 3:

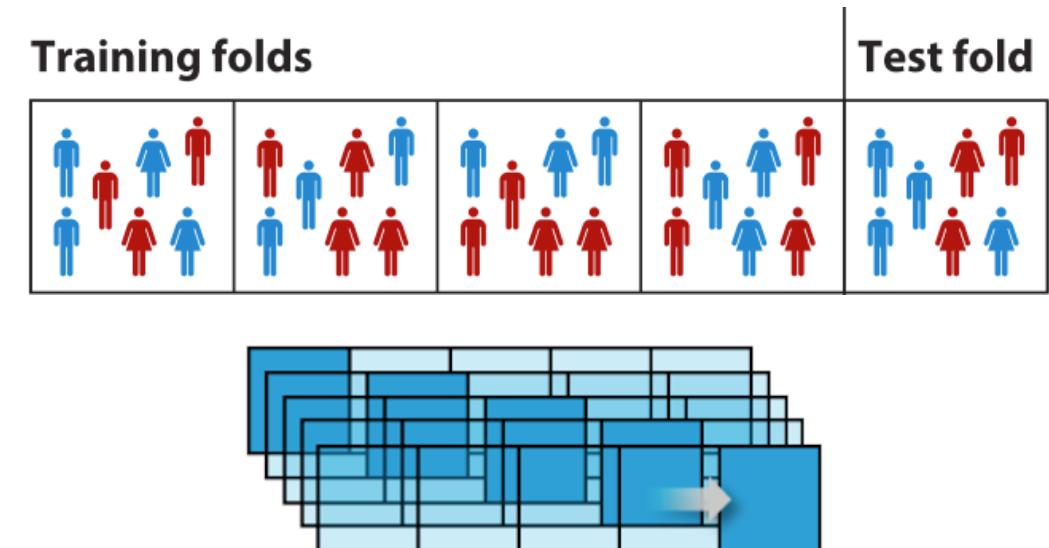
- **Prüfung der Generalisierbarkeit:** Um zu prüfen, ob das trainierte Modell gut funktioniert, müssen wir es validieren
- **Validierung:** Modell an Fällen erproben die nicht im Trainingsset dabei waren (d.h. an bis dahin unbekannten Personen) → entweder man erhebt also neue Daten, oder man teilt einen vorhandenen Datensatz in ein Trainings- und ein Testset
- **VORSICHT:** strikte Trennung zwischen Trainings- und Testdaten (Data-Leakage) → sonst droht ein **Overfit** (Überanpassung an die Daten)
 - Unfairer Vorteil: Vorhersage von Fällen, die das Modell schon "kennt"
 - Modell würde sonst besser performen, als in der Realität

Klassen von Machine Learning

Logik der Prädiktiven Modellierung

Wie trenne ich einen vorhandenen Datensatz in ein Trainings- und Testset?

- Einfachstes Szenario: **Simple Split** → Datensatz wird z.B. 50/50 in Trainings- und Testdatensatz geteilt
 - Problem: Forscher könnte genau so trennen, dass Modell gut funktioniert (das wäre Schummeln)
- Lösung: **(k-fold) Nested-Cross-Validation**
 - Datensatz wird in k Folds geteilt (hier k=5)
 - Runde 1: Training: 1-4 Folds, Test: 5. Fold
 - Runde 2: Training: 1,2,3,5 Folds, Test: 4. Fold
 - Runde 3: Training: 1,2,4,5 Folds, Test: 3. Fold
 - ...
 - Am Ende Genauigkeitsmaße ausrechnen
- Vorteil: Kein Overfit + ganzer Datensatz kann ohne Data-Leakage genutzt werden

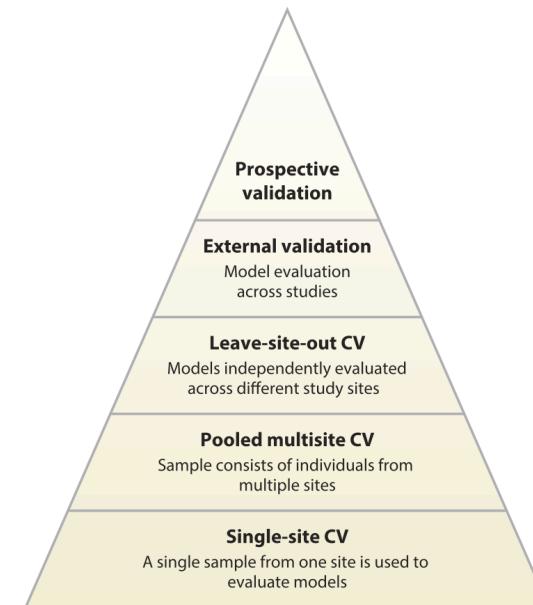


Klassen von Machine Learning

Logik der Prädiktiven Modellierung

Generalisierbarkeitspyramide:

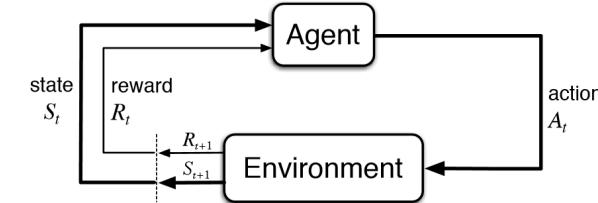
- Je unterschiedlicher die Kontexte (z.B. zeitlich, geografisch, kulturell) von Training und Test Folds, desto generalisierbarer ist das Modell
- Leave-Site-out → Folds = einzelne Studienzentren: Modell funktioniert nicht nur in Klinik A sondern auch in Klinik B
- External → Folds = einzelne Studien: Modell funktioniert nicht nur in Studie A sondern auch in Studie B
- Prospective → Folds = Zeit: Modell kann Fälle vorhersagen, die in der Zukunft liegen



Klassen von Machine Learning

Reinforcement Learning

- Beinhaltet das Erlernen von Funktionen und Entscheidungsregeln, um erhaltene Belohnungen zu maximieren (Fehler zu minimieren)
- Output dieser Algorithmen kann ein komplexes Verhalten sein
- Kann implizit auch Klassifikationen oder Gruppierungen beinhalten
- Beispiele sind Schach spielen lernen, autonomes Fahren, etc.



- Output: z.B. Lenkrad links/rechts drehen
- Objective Function: schnell ans Ziel ohne Unfall/Risiko

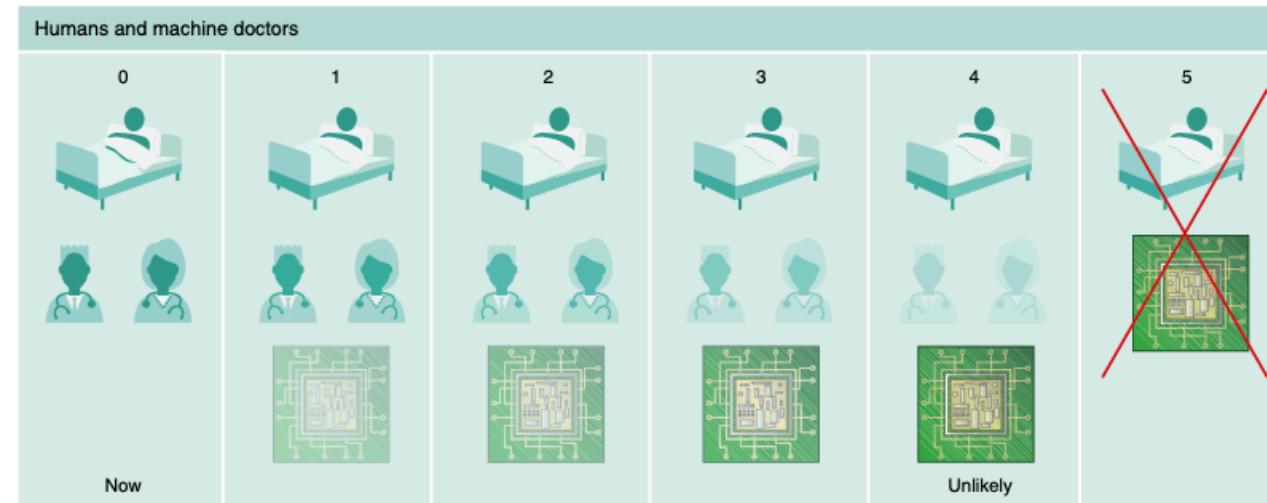
Mensch-Maschine Interaktion

Analogie: Stufen autonomen Fahrens

Stufe	Name	Interaktion
0	No automation	Fehlen von Assistenzsystemen.
1	Driver assistance	Systeme, die dem Fahrer helfen, die Geschwindigkeit zu halten oder in der Spur zu bleiben, aber dem Fahrer die Kontrolle überlassen.
2	Partial automation	Die Kombination aus automatischer Geschwindigkeits und Lenkung - zum Beispiel Tempomat und Spurhaltung.
3	Conditional automation	Automatisierte Systeme, die die Umgebung steuern und überwachen, aber auf einen menschlichen Fahrer als Unterstützung angewiesen sind.
4	High automation	Automatisierte Systeme, die alles erledigen - ohne menschliche Unterstützung - aber nur unter bestimmten Umständen.
5	Full automation	Der echte elektronische Chauffeur: Er behält die volle Kontrolle über das Fahrzeug, braucht keine menschliche Unterstützung und fährt unter allen Bedingungen.

Big Data und Machine-Learning

Mensch-Maschine Interaktion



- Stufe 5, d. h. eine vollständige Automatisierung ohne die Möglichkeit der Unterstützung von Klinikpersonal ist nicht das Ziel.
- Auch Stufe 4 mit menschlicher Unterstützung unter sehr begrenzten Bedingungen ist nicht das Ziel.
- Angestrebt werden Synergieeffekte, d. h. die Kombination von Funktionen, die Maschinen am besten erledigen, mit solchen, die am besten für Kliniker:innen geeignet sind.

Klinische Akzeptanz von KI (Koutsouleris, 2021)

Anforderungen:

- Medizinische KI muss mindestens so genau sein wie menschliche Heuristik
- KI muss mindestens so generalisiert sein wie klinische Entscheidungsfindungen
- KI muss so trainiert werden, dass sie zu einer effektiven Nutzung bestehender Ressourcen führt
- Die Anwendungsindikation von medizinischer KI muss klar definiert ist
- Der Bias von KI muss offen gelegt werden können
- KI muss transparent bezüglich der eigenen Unsicherheit sein, um übermäßiger Technikgläubigkeit vorzubeugen
- KI muss handlungsrelevante Informationen produzieren
- KI muss klinisch kontrolliert werden können (was eine entsprechende Ausbildung im Studium erfordert)

Take-Aways

- **Ambulatory Assessment (AA)** ist die wiederholte Erfassung von psychologischer (Kontext-)Variablen in natürlicher Umgebung.
- Der Zeitpunkt einer ESM/EMA Erhebung kann gemäß dem **Signal-contingent sampling** oder **Event-contingent sampling** vorgegeben werden.
- **Ecological Momentary Interventions (EMI)** sind alltagsnahe, auf den Moment/Tag bezogene Interventionen
- **Machine Learning:** Computerfunktionalität, die nicht explizit programmiert wurde, sondern selbstständig "lernt".
- **Unsupervised ML** → Mustererkennung, **Supervised ML** → Prädiktion, **Reinforcement Learning** → Operation
- Vorhersagegenauigkeit (**Accuracy**, **Sensitivität** und **Spezifität**) eines Prädiktionsmodells muss mittels **Validierung** auf neuen Daten geprüft werden.
- **Overfit:** Überanpassung an die Daten → führt zu verringelter Generalisierbarkeit.

Schlüssel-/Fachbegriffe der heutigen Vorlesung

Ambulatory Assessment	Fixed scheme	Unsupervised ML	Generalisierbarkeit
Experience Sampling Methode (ESM)	Random scheme	Supervised ML / Prädiktive Modellierung	Overfit
Ecological Momentary Assessment (EMA)	Ecological Momentary Interventions (EMI)	Reinforcement/Reward Learning	Kreuzvalidierung (CV)
Mobile bzw. Ambient Sensing	Kapazitätsproblem	Trainings- und Testdatensatz	Simple Split
Time-/Signal-contingent sampling	Big Data	Accuracy	(k-fold) Nested-Cross-Validation
Event-contingent sampling	Machine Learning (ML)	Sensitivität	Generalisierbarkeitspyramide
	Perzeptron	Spezifität	Leave-Site-Out CV
	Präzisionsmedizin (taylored treatments)		External validation
			Prospective CV

[zurück zur heutigen Übersicht der Vorlesung →](#)
[zum Quiz zur Wissensprüfung →](#)

Übernächste Woche: Zusammenfassung + Fragestunde

Damit die Zeit in dieser Stunde gut genutzt werden kann, würde ich Sie bitten Ihre Fragen zur Vorlesung in folgenden Dokument zu sammeln: <https://docs.google.com/document/d/1tWKYDSsRBc7MxoYnH4Cm6tyCBMkouVYJc0NMcGWDKFY/edit>

Je nach Fragen-Aufkommen werde ich dann Fragen priorisieren und versuchen daraus eine interaktive Lern-Möglichkeit für alle zu schaffen.

Ich habe ChatGPT um ein Bild für die Fragestunde gebeten. Hier das Ergebnis:

