

# Grundlagen der Diagnostik

Sitzung 10

## Diagnostik im digitalen Zeitalter



We are happy to share our materials openly:

The content of these Open Educational Resources by Lehrstuhl für Psychologische Methodenlehre und Diagnostik, Ludwig-Maximilians-Universität München is licensed under CC BY-SA 4.0. The CC Attribution-ShareAlike 4.0 International license means that you can reuse or transform the content of our materials for any purpose as long as you cite our original materials and share your derivatives under the same license.

# Organisatorisches zur Klausur

- Fragestunde: **Montag, 03.08, 14:15-15:45** ➔ Zoom-Link im Moodle!
- Modulprüfung: **Mittwoch, 05.08., 12:00-14:00 Uhr**
  - Seien Sie bitte pünktlich!
  - Prüfungsräume und Raumaufteilung:
    - Nachnamen mit Anfangsbuchstaben A bis H: **Raum B001, Oettingenstr. 67**
    - Nachnamen mit Anfangsbuchstaben I bis Q: **Raum 2U01, Leopoldstr. 13**
    - Nachnahmen mit Anfangsbuchstaben R bis Z: **Raum B138, Theresienstraße 39**
  - Erlaubtes Hilfsmittel: Taschenrechner
  - Achtung: Smartwatches sind während der Klausur nicht erlaubt!
  - Achtung: Verwenden Sie bitte dokumentenechte Stifte (d.h., kein Bleistift)!

- Was ist klausurrelevant? → **Siehe Syllabus!**
  - „Alle Foliensätze aus der Vorlesung und die Übungsblätter aus dem UK sind klausurrelevant. Die einzige Ausnahme bilden einzelne Folien, die **mit einem roten Punkt** im oberen rechten Eck gekennzeichnet sind – diese sind nicht klausurrelevant.
  - Die in den Folien und Übungsblättern angegebenen **Formeln werden in der Klausur angegeben** und müssen nicht auswendig gelernt werden.“ → **Achtung: Bei den Gleichungen im Foliensatz 07 zu „Urteilen und Fehlern“ handelt es sich um Definitionen und diese werden nicht angegeben. Auch die Berechnung von Normwerten in Foliensatz 08 sollten Sie beherrschen.**
- Wie sehen die Klausuraufgaben aus? → **Siehe Syllabus!**
  - „Die Klausur wird etwa 40 Aussagen enthalten, die als „wahr“ oder „falsch“ beurteilt werden müssen.“
  - „Die Aussagen sind jeweils in Blöcken zu Aufgaben zusammengefasst, die sich entweder auf ein gemeinsames Thema aus der Lehrveranstaltung oder auf eine gemeinsame Aufgabenstellung (eventuell mit R-Output) beziehen.“
  - „Jede einzelne Aussage in der Klausur wird immer entweder mit 1 Punkt (Aussage korrekt beurteilt) oder mit 0 Punkten (Aussage nicht korrekt beurteilt) gewertet. Es gibt keinen Punktabzug für falsche Antworten innerhalb einer Aufgabe. Sie können somit bei z.B. 40 Aussagen maximal 40 Punkte (= alles richtig) und minimal 0 Punkte (= alles falsch) erzielen.“

- Wie ist der Antwortbogen auszufüllen? → **Siehe Altklausur im Moodle!**
  - Tragen Sie unbedingt Ihren Namen und Ihre Matrikelnummer ein!
  - Übertragen Sie Ihre Matrikelnummer in das darunterliegende Kästchen-Schema!
  - Setzen Sie bei jeder Aufgabe **für jede richtige Aussage ein Kreuz** in das entsprechende Kästchen. Für jede falsche Aussage lassen Sie das entsprechende Kästchen frei. Es können keine bis alle Aussagen pro Aufgabe richtig sein.
  - Achten Sie darauf, dass **eindeutig** erkennbar ist, in welchen Kästchen Sie Ihre Kreuze gesetzt haben!
- Wie geht es nach der Klausur weiter? → **Siehe Syllabus!**
  - Die Klausur wird samt Musterlösung (richtig/falsch Lösungen) zeitnah nach dem Klausurtermin im Vorlesungs-Moodle hochgeladen
  - Wir bemühen uns, die Klausur schnell zu korrigieren, und benachrichtigen Sie mittels einer Ankündigung im Vorlesungs-Moodle, sobald die Noten im LSF eingetragen sind

| Sitzung | Datum    | Thema                          | Themenblock                                                                                            |
|---------|----------|--------------------------------|--------------------------------------------------------------------------------------------------------|
| 1       | 16.04.26 | Einführung I                   | Definitionen; diagnostischer Prozess; gesetzlicher Rahmen; diagnostische Entscheidungen; Gütekriterien |
| 2       | 23.04.26 | Einführung II                  |                                                                                                        |
| 3       | 30.04.26 | Einführung II fertig           |                                                                                                        |
| 4       | 07.05.26 | Verhaltensbeobachtung          | Verhaltensbeobachtung als diagnostisches Verfahren                                                     |
| /       | 14.05.26 | entfällt wegen Feiertag        |                                                                                                        |
| 5       | 21.05.26 | Beobachterübereinstimmung I    | Maße zur Bestimmung der Übereinstimmung zwischen verschiedenen Beobachtern/Ratern                      |
| 6       | 28.05.26 | Beobachterübereinstimmung II   |                                                                                                        |
| /       | 04.06.26 | <i>entfällt wegen Feiertag</i> |                                                                                                        |
| 7       | 11.06.26 | Interviews                     | Interviews als diagnostisches Verfahren                                                                |
| 8       | 18.06.26 | Urteile und Fehler             | Diagnostische Urteilsbildung und Güte von Urteilen                                                     |
| 9       | 25.06.26 | Einzelfalldiagnostik I         | Methoden der Einzelfalldiagnostik aus der frequentistischen und bayesianischen Statistik               |
| 10      | 02.07.26 | Einzelfalldiagnostik II        |                                                                                                        |
| 11      | 09.07.26 | Digitale Diagnostik            | Diagn. Einsatz von digitalen Daten und maschinellem Lernen                                             |

- ➔ In der heutigen Vorlesung beschäftigen wir uns mit dem Einsatz digitaler Daten und maschineller Lernverfahren in der psychologischen Diagnostik

# Diagnostische Verfahren im digitalen Zeitalter

Klassische Instrumente der psychologischen Diagnostik werden im digitalen Zeitalter neu interpretiert und/oder erweitert...

- Psychologische Tests und Fragebögen
- Interviews
- Verhaltensbeobachtung

Klassische Instrumente der psychologischen Diagnostik werden im digitalen Zeitalter neu interpretiert und/oder erweitert...

- **Psychologische Tests und Fragebögen**

- Online- und App-basierte Testverfahren, Durchführung auf mobilen Endgeräten (z.B. Tablets, Smartphones)
  - ➔ orts- und zeitunabhängige Durchführung
  - ➔ wiederholte Messungen leichter umsetzbar
  - ➔ automatische Auswertung und Rückmeldung
- Beispiele:
  - Mobiler Intelligenztest fürs Smartphone (Schuhfried, 2022)
  - Ecological Momentary Assessments (EMAs) zur wiederholten Erfassung von momentanen Gedanken/Gefühlen/Verhaltensweisen in Echtzeit durch Ultra-Kurz-Fragebögen, z.B. am Smartphone (s., Wrzus & Neubauer, 2023)

Klassische Instrumente der psychologischen Diagnostik werden im digitalen Zeitalter neu interpretiert und/oder erweitert...

- **Interviews**

- Videointerviews, KI-gestützte Chatbots
  - ➔ größere räumliche (und zeitliche) Flexibilität
  - ➔ höheres Maß der Standardisierung (bei Chatbots)
  - ➔ automatische Transkription und Ratings mittels KI
  - ➔ zusätzliche Analyse sprachlicher Merkmale
- Beispiele:
  - Auswahlinterviews via Zoom (z.B. Oliffe et al., 2021)
  - Strukturierte Klinische Interviews per Chatbot (siehe Bendig et al., 2022; He et al., 2023)

Klassische Instrumente der psychologischen Diagnostik werden im digitalen Zeitalter neu interpretiert und/oder erweitert...

- **Verhaltensbeobachtung**
  - Videoaufzeichnungen, digitale Verhaltensdaten (siehe auch Folie 16)
    - ➔ automatische Aufzeichnung und Ratings von Daten
    - ➔ Zugriff auf diverse Verhaltensparameter
    - ➔ „Beobachtung“ im natürlichen Alltag
  - Beispiele:
    - Verbale und paraverbale Variablen im AC (z.B. Grunenberg et al., 2026)
    - Digital Footprints in auf Social Media (z.B. Goldbeck et al., 2011; Kosinski et al., 2013)

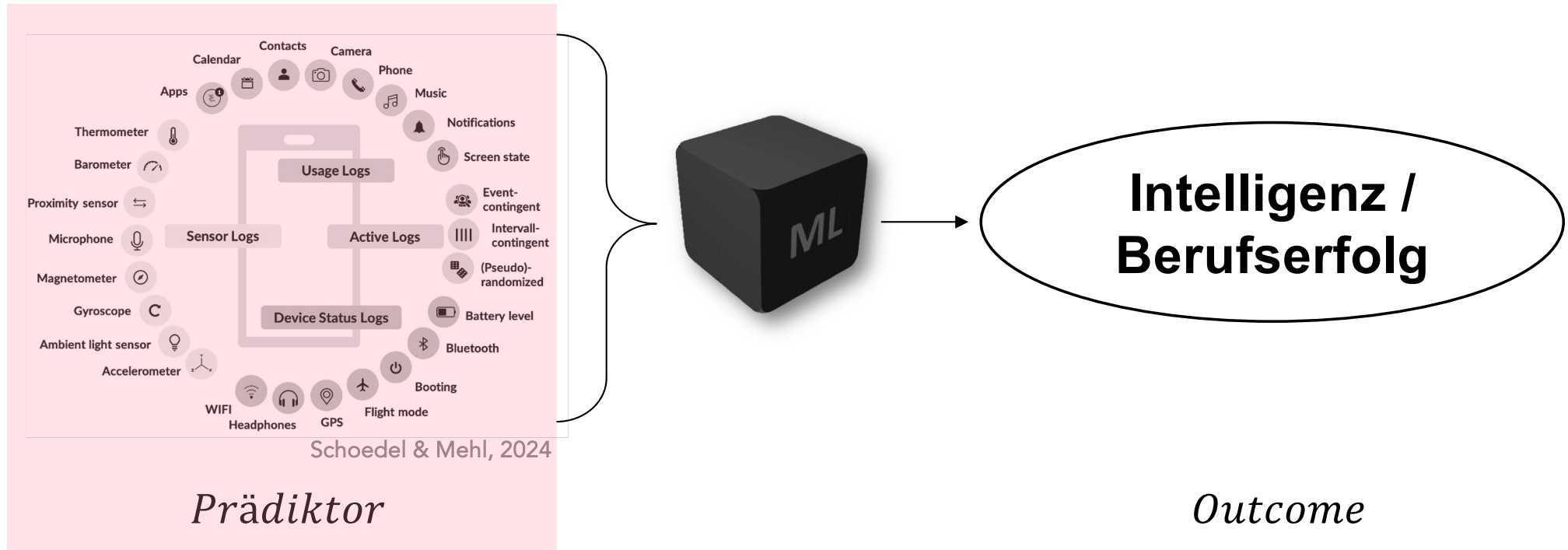
# Automatisierte Erfassung psychologischer Merkmale

## Klassische Diagnostik:

- Erfassung psychologischer Merkmale (z.B. Intelligenz, Depressivität) mittels Tests, Interviews und Verhaltensbeobachtung
- Ableitung diagnostischer Urteile (z.B. Eignung; Diagnose) aus den gewonnenen Informationen

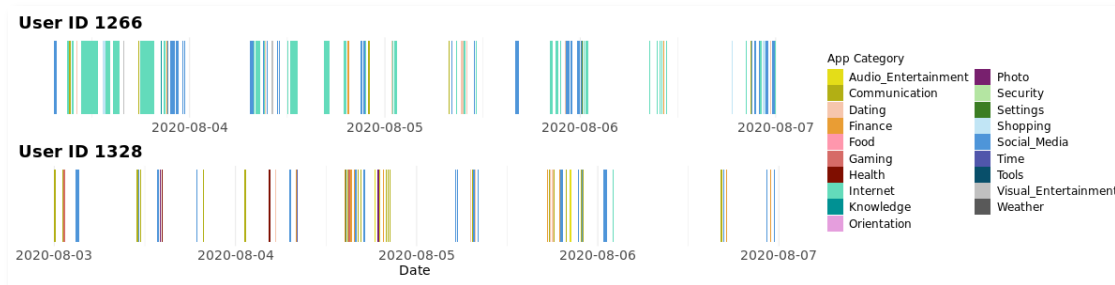
## Automatisierte Diagnostik:

- **Digitale Verhaltensdaten** als Informationsquelle
- **Machine Learning** zur Vorhersage psychologischer Merkmale (z.B. Intelligenz) oder Kriterien (z.B. Berufserfolg)

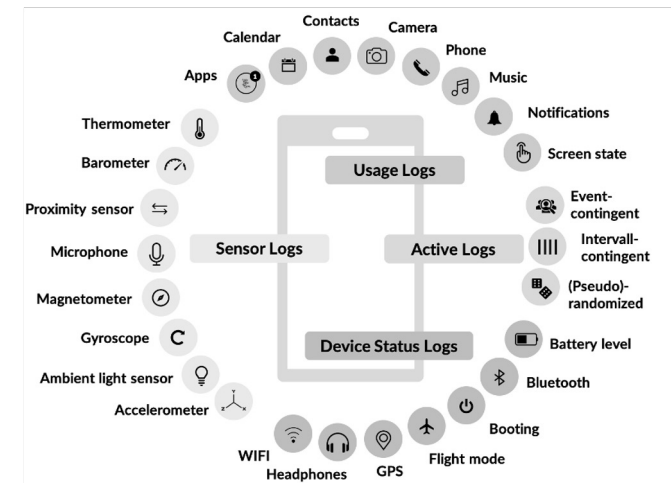


Wir geben digitale Verhaltensdaten in ein Machine-Learning-Modell um die Ausprägung eines psychologischen Merkmals (z.B. Intelligenz) oder eines Kriteriums (z.B. Berufserfolg) vorherzusagen

- Die Nutzung digitaler Geräte (z.B. Smartphones, Wearables) und Dienste (z.B. Google, Instagram) hinterlässt **digitale Verhaltensspuren**
  - z.B. Nutzungsdaten, Sprachdaten, Sensordaten
- Diese Daten können über spezielle Apps (= **Mobile Sensing**) erfasst oder von Plattformanbietern bereitgestellt werden
- Diese digitalen Spuren erlauben Rückschlüsse auf das **Verhalten im Alltag**, z.B. am Smartphone:
  - Bildschirm an/aus → Nutzungsdauer, Schlafrhythmus
  - WhatsApp-Nutzung → Kommunikationsfrequenz
  - GPS-Daten → Aufenthaltsorte, Mobilitätsradius



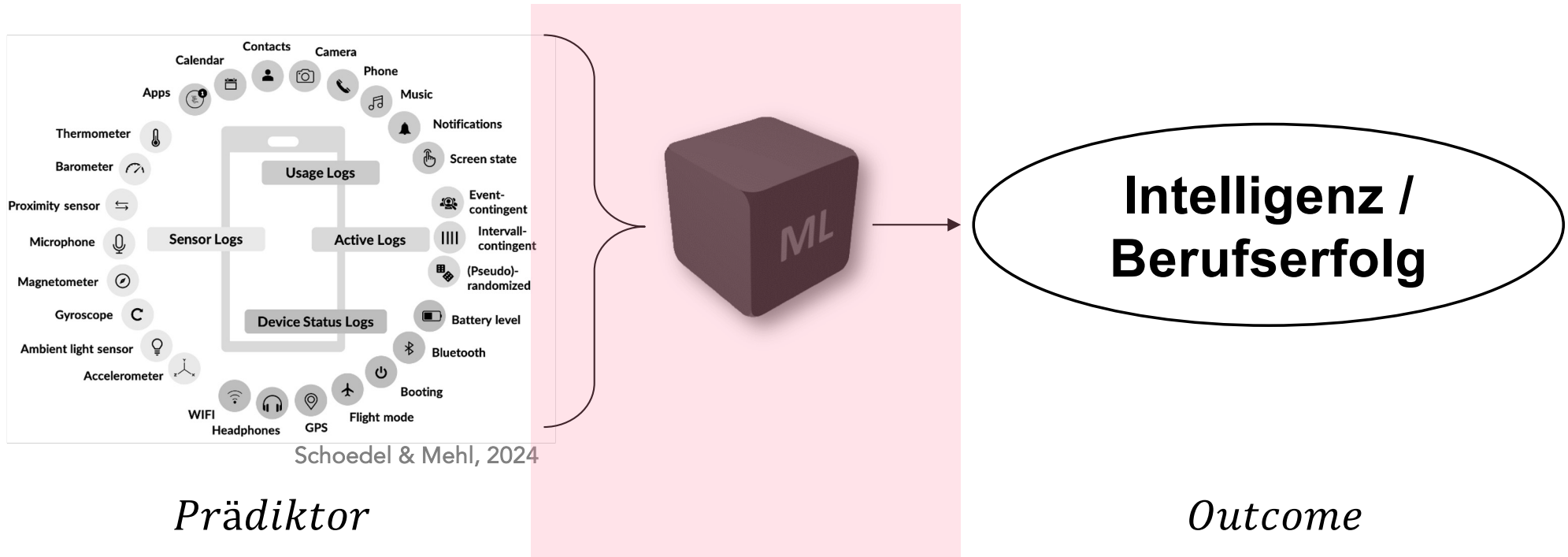
Schoedel et al., 2026



Schoedel & Mehl, 2024

## Vorteile digitaler Verhaltensdaten

- **Umfangreiche und objektive Datenerfassung**
  - Automatische Datenerfassung ohne großen Aufwand für die Zielperson
  - Integration von hunderten bis tausenden Verhaltensmerkmalen
  - Geringere Anfälligkeit für Verzerrungen als Selbstberichte (z.B. soziale Erwünschtheit oder Erinnerungslücken)
- **Verhalten im Alltag statt im Labor**
  - Hohe ökologische Validität durch Erfassung im natürlichen Lebensumfeld
  - Kontinuierliche Datenerfassung über längere Zeiträume



Wir geben digitale Verhaltensdaten in ein Machine-Learning-Modell um die Ausprägung eines psychologischen Merkmals (z.B. Intelligenz) oder eines Kriteriums (z.B. Berufserfolg) vorherzusagen

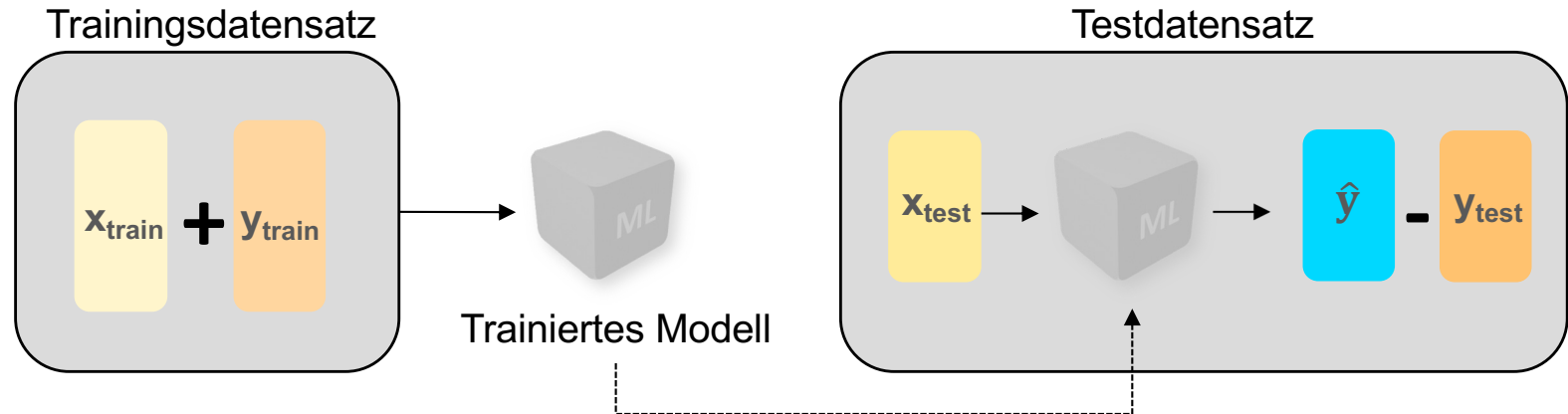
- Machine Learning lernt anhand vorhandener Daten Zusammenhänge zwischen Prädiktoren und einem Outcome, um Vorhersagen für neue Personen zu treffen
  - ➔ Ziel ist die **möglichst genaue Vorhersage** des Outcomes, nicht die Messung eines Merkmals anhand eines psychologischen Messmodells
    - z.B. „Welche verfügbaren Informationen erlauben die beste Vorhersage von Intelligenz?“ statt „Wie messen wir Intelligenz?“
- Hierzu wird ein Modell anhand eines großen Datensatzes **trainiert** (d.h., die Modellparameter werden geschätzt) und anschließend **an unabhängigen Daten überprüft**, um seine Generalisierbarkeit zu bewerten
- Beim sog. „**Supervised Learning**“ sind sowohl Prädiktoren als auch Outcome während des Trainings des Modells bekannt; später sagt das Modell das Outcome **alleine anhand der Prädiktoren** vorher
  - Beispiel: Ein Modell wird mit Smartphone-Nutzungsdaten und Intelligenztestwerten trainiert und sagt anschließend die Intelligenz neuer Personen anhand ihrer Smartphone-Nutzung vorher
- Die konkrete Modellform ist dabei zweitrangig – sie kann von einer linearen Regression bis zu komplexen non-linearen Modellen reichen (siehe Master-VL)



| x   | y   |
|-----|-----|
| 5   | 2   |
| 6   | 1   |
| 1   | 2   |
| 2   | 0   |
| 5   | 5   |
| 2   | 1   |
| ... | ... |

Training

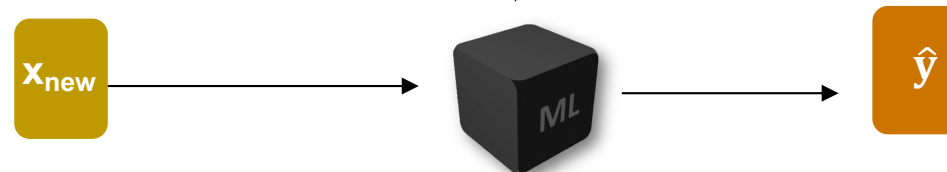
Test



## Training und Evaluation des Modells



Wir schauen uns anhand des Testdatensatzes an, wie gut die Vorhersage funktioniert hat. Dies liefert uns eine **Schätzung** dafür, wie gut das **Modell, das auf dem Gesamtdatensatz trainiert** wurde, später auf neuen Daten funktioniert.



## Anwendung des Modells auf neuen Daten

## Ziele automatisierter Diagnostik

Das Hauptziel vieler diagnostischer Fragestellungen besteht darin, basierend auf den verfügbaren Informationen (Prädiktoren) **möglichst präzise Entscheidungen oder Vorhersagen hinsichtlich eines interessierenden Kriteriums** abzuleiten.

Machine Learning wird hierbei auf zwei Weisen eingesetzt:

- Vorhersage psychologischer Merkmale (z.B. Intelligenz, Persönlichkeit, Depressivität) → Ziel ein klassisches Instrument der Diagnostik (z.B. Test, Interview) zu ersetzen, um **diagnostische Daten zu gewinnen**
- Vorhersage diagnostisch relevanter Kriterien (z.B. Berufserfolg, Onset einer Depression, Suizidversuch) → Ziel direkt eine **diagnostische Entscheidung zu fällen**
  - *Bsp. Personalpsychologie:* Ziel von Unternehmen ist es, möglichst geeignete Bewerberinnen zu finden → Vorhersage von Berufserfolg
  - *Bsp. Klinische Psychologie:* Ziel von Therapeutinnen ist es möglichst frühzeitig kritische Symptomverläufe zu erkennen und präventiv zu handeln → Vorhersage eines Suizidversuchs

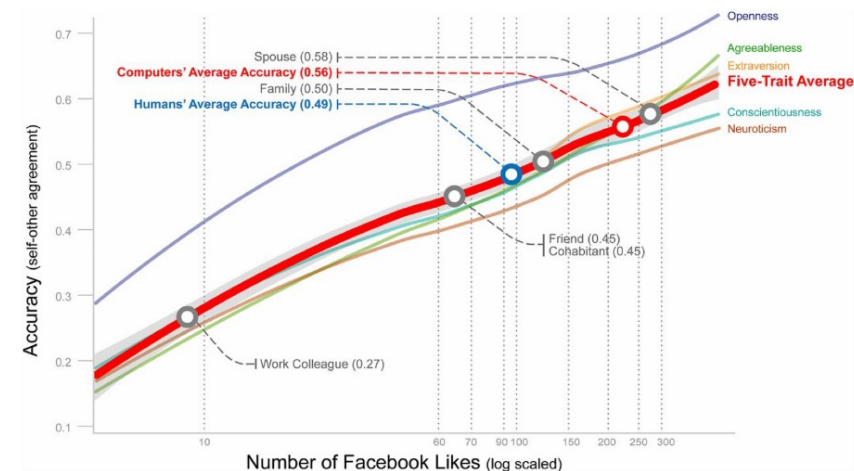


## Beispiel 1

### Facebook-Likes ➔ Persönlichkeit

- $N = 70.520$
- Prädiktoren: Facebook Likes
- Kriterium: Big Five Dimensionen
- Durchschnittliche Korrelation zwischen vorhergesagtem Persönlichkeitswert und selbst-berichtetem Persönlichkeitswert:  $r = .56$

Computer-based personality judgment accuracy (y axis), plotted against the number of Likes available for prediction (x axis).



Wu Youyou et al. PNAS 2015;112:4:1036-1040

©2015 by National Academy of Sciences

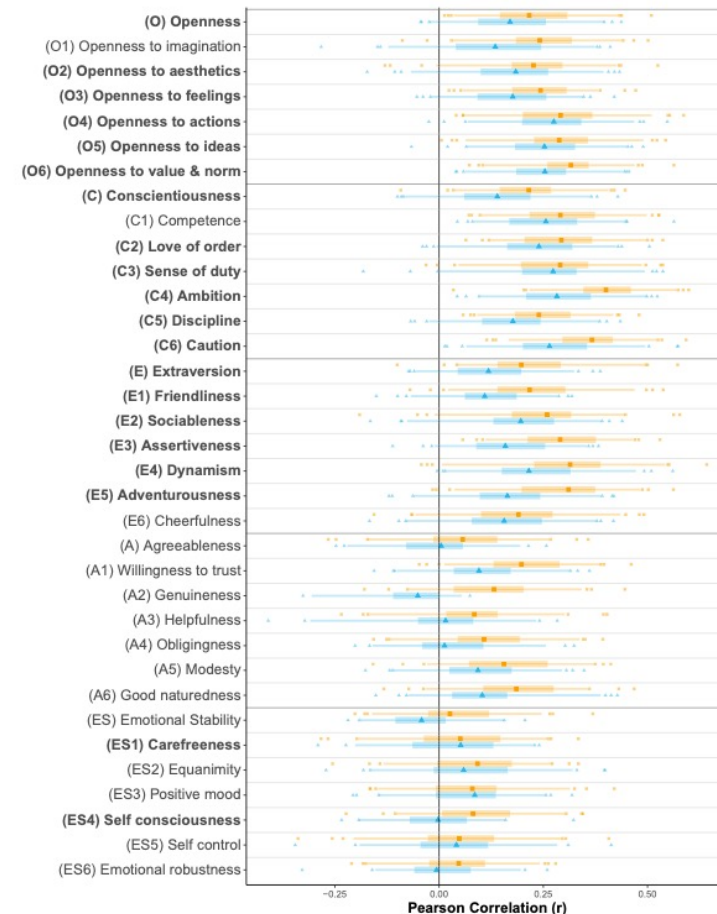
PNAS



## Beispiel 2

### Smartphone-Nutzung → Persönlichkeit

- $N = 624$ , 30 Tage
- Prädiktoren: ca. 16.000 Variablen zu Mobilität, Kommunikation, App-Nutzung, Smartphone-Aktivität, Tag-Nacht-Aktivität
- Kriterium: Big Five Dimensionen & Facetten
- Durchschnittliche Korrelation zwischen vorhergesagtem Persönlichkeitswert und selbst-berichtetem Persönlichkeitswert:  $r = .40$

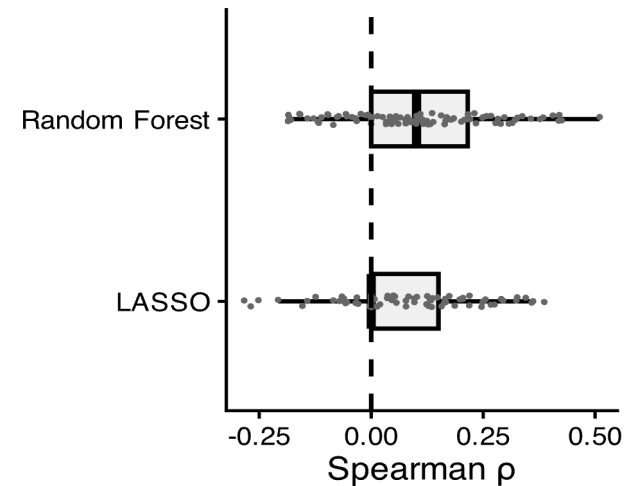




## Beispiel 3

### Musikhörverhalten → Intelligenz

- $N = 185$ , 5 Monate
- Prädiktoren: 215 Variablen zu Musikhörverhalten und -präferenzen
- Kriterium: General Cognitive Functioning
- Durchschnittliche Korrelation zwischen vorhergesagtem Intelligenzwert und getestetem Intelligenzwert:  $r = .10$



- Die Vorhersagegüte psychologischer Merkmale und Kriterien ist derzeit sehr unterschiedlich und hängt stark von der jeweiligen Anwendung ab
- Die Modellgüte wird u. a. beeinflusst durch
  - die Qualität und Größe der Trainingsdaten,
  - das verwendete Machine-Learning-Modell
- Die Vorhersage psychologischer Merkmale wird zusätzlich dadurch erschwer, dass
  - viele psychologische **Zusammenhänge klein bis moderat** sind
  - digitale Daten nur **äußeres Verhalten**, nicht jedoch inneres Erleben, erfassen
- Die Forschung in diesem Bereich entwickelt sich derzeit jedoch rasant!
- **Digitale Diagnostik kann klassische diagnostische Verfahren derzeit sinnvoll ergänzen, ersetzt sie in der Einzelfalldiagnostik aber noch nicht vollständig**



# Herausforderungen digitaler Diagnostik



## Ethik, Datenschutz & Transparenz

- Digitale Verhaltensdaten sollten ausschließlich für **ethisch vertretbare Zwecke** eingesetzt werden!
  - Z.B. kann die Vorhersage psychischer Erkrankungen der Prävention dienen, sollte jedoch nicht zur Benachteiligung Betroffener (z.B. bei Krankenkassenbeiträgen) genutzt werden.
- Digitale Verhaltensdaten sind häufig personenbezogen und erfordern einen **datenschutzkonformen** Umgang!
- Automatisierte diagnostische Entscheidungen (sog. „Profiling“) erfordern laut DSGVO eine **ausdrückliche Einwilligung** der Betroffenen
  - Daraus ergibt sich eine **Transparenzpflicht**: Betroffene müssen vollumfänglich und verständlich über die Logik und Auswirkungen des Profilings informiert werden
  - Aber: Können Betroffene bei der hohen Komplexität der eingesetzten Verfahren tatsächlich „informiert“ einwilligen?



## Interpretierbarkeit

- Im Machine Learning werden häufig Modelle mit der höchsten Vorhersagegenauigkeit bevorzugt
- Diese sind jedoch oft komplex und nur **eingeschränkt interpretierbar**: Bei sog. „**Black Box**“-**Modellen** lässt sich nicht direkt erkennen, welche Prädiktoren die Vorhersage beeinflussen und in welche Richtung
  - Gegenbeispiel: Ein einfaches, gut interpretierbares Modell wäre z.B. die Multiple Lineare Regression mit ihren Regressionsgewichten (z.B. längere Social Media-Nutzung → höhere vorhergesagte Extraversion)
- Dies steht teilweise im Gegensatz zum klassischen Ziel der Psychologie, Verhalten zu erklären und theoretische Zusammenhänge zu verstehen
- Daher sollten Methoden der **Explainable AI (xAI)** eingesetzt werden, um zu untersuchen, welche Prädiktoren ein Modell tatsächlich für seine Vorhersagen nutzt



## Interpretierbarkeit

Verwendet das Modell die „richtigen“ Prädiktoren? Ein Beispiel...



Predicted: Wolf  
True: Wolf



Predicted: husky  
True: husky



Predicted: Wolf  
True: Wolf



Predicted: Wolf  
True: Wolf



Predicted: husky  
True: husky



Predicted: Wolf  
True: Husky

[https://www.sintef.no/contentassets/ade1a076009c4dbb9d5e9a7fcaebfed4/1\\_sintef\\_hva-er-kunstig-intelligens.pdf](https://www.sintef.no/contentassets/ade1a076009c4dbb9d5e9a7fcaebfed4/1_sintef_hva-er-kunstig-intelligens.pdf)



## Fairness

- Ein Verständnis dafür, wie ein Modell seine Vorhersagen macht, spielt auch eine Rolle bei der Beurteilung der **Fairness** des Modells!
- In einem fairen Modell sollten die Vorhersagen nicht durch irrelevante oder geschützte Personenmerkmale beeinflusst werden und keine Personengruppen (z.B. aufgrund von Geschlecht, Alter oder ethnischer Herkunft) **systematisch benachteiligt** werden.
- Beispiel: Algorithmus zur Personalauswahl bei Amazon (Dastin, 2018)
  - Modell zur Vorhersage von Berufserfolg basierend auf Lebenslauf-Texten
  - Das Modell bevorzugte Kandidaten, die sich selbst mit Wörtern beschrieben, die eher in Lebensläufen von männlichen Ingenieuren zu finden waren
  - Anwendung des Modells führt zur Einstellung eines größeren Anteils an Männern → Geschlechtsdiskriminierung



## Generalisierbarkeit

- Machine-Learning-Modelle lernen Zusammenhänge aus **einer bestimmten Stichprobe**, die nicht zwangsläufig auf **andere Populationen** (z.B. andere Altersgruppen oder Länder) oder **Anwendungskontexte** (z.B. klinische Patienten statt gesunder Studierender) übertragbar sind
- Daher ist eine **unabhängige Validierung** der Modelle in neuen Stichproben und Anwendungskontexten unerlässlich
- Zudem können sich Zusammenhänge über die Zeit verändern (**Konzept-Drift**), sodass Vorhersagemodelle regelmäßig überprüft und aktualisiert werden müssen
  - Beispiel: Die Nutzung von SMS war früher ein relevanter Prädiktor für Extraversion, spielt jedoch heute kaum noch eine Rolle (Stachl, Pargent et al., 2020)



- Die Digitalisierung eröffnet vielfältige neue Möglichkeiten für die psychologische Diagnostik – von der kontinuierlichen Erfassung von Verhalten bis hin zur automatisierten Vorhersage psychologischer Merkmale.
- Gleichzeitig bestehen weiterhin methodische, ethische und rechtliche Herausforderungen, die vor einem breiten Einsatz in der Praxis sorgfältig berücksichtigt werden müssen.
- Psychologinnen sind gefordert, die Entwicklung digitaler Diagnostik kritisch zu begleiten und ihre Qualität fundiert beurteilen zu können. Dies setzt voraus, neue Verfahren und wissenschaftliche Erkenntnisse kontinuierlich zu verfolgen.

## Das war's schon fast für heute...

- *Ausblick:* Denken Sie an die Fragestunde am **Montag, 03.08, 14:15-15:45!**
- Aber zuerst: Gibt es offene Fragen zur heutigen Vorlesung?

- Bendig, E., Erb, B., Schulze-Thuesing, L., & Baumeister, H. (2022). The next generation: chatbots in clinical psychology and psychotherapy to foster mental health—a scoping review. *Verhaltenstherapie*, 32(Suppl. 1), 64-76.  
<https://doi.org/10.1159/000501812>
- Dastin, J. (2018). Amazon scraps secret AI recruiting tool that showed bias against women. Towards Data Science. Retrieved from <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G>
- Golbeck, J., Robles, C., Edmondson, M., & Turner, K. (2011, October). Predicting personality from twitter. In 2011 IEEE third international conference on privacy, security, risk and trust and 2011 IEEE third international conference on social computing(pp. 149-156). IEEE. <https://doi.org/10.1109/PASSAT/SocialCom.2011.33>
- Harari, G. M., Lane, N. D., Wang, R., Crosier, B. S., Campbell, A. T., & Gosling, S. D. (2016). Using smartphones to collect behavioral data in psychological science. *Perspectives on Psychological Science*, 11 (6), 838–854.  
[doi:10.1177/1745691616650285](https://doi.org/10.1177/1745691616650285)
- He, Y., Yang, L., Qian, C., Li, T., Su, Z., Zhang, Q., & Hou, X. (2023). Conversational agent interventions for mental health problems: systematic review and meta-analysis of randomized controlled trials. *Journal of medical Internet research*, 25, e43862. <https://doi.org/10.2196/43862>
- Kosinski, M., Stillwell, D., & Graepel, T. (2013). Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the national academy of sciences*, 110(15), 5802-5805. <https://doi.org/10.1073/pnas.1218772110>
- Oliffe, J. L., Kelly, M. T., Gonzalez Montaner, G., & Yu Ko, W. F. (2021). Zoom interviews: Benefits and concessions. *International journal of qualitative methods*, 20, 16094069211053522.  
<https://doi.org/10.1177/16094069211053522>
- Pargent, F., Schoedel, R., & Stachl, C. (2023). Best practices in supervised machine learning: A tutorial for psychologists. *Advances in Methods and Practices in Psychological Science*, 6(3), 25152459231162559.  
<https://doi.org/10.1177/25152459231162559>
- Rust, J., Kosinski, M., & Stillwell, D. (2020). *Modern Psychometrics: The Science of Psychological Assessment* (4th ed.). Routledge. <https://doi.org/10.4324/9781315637686>

- Schoedel, R., Altgassen, E., Back, M. D., Buss, M., Bühner, M., Etzel, J. M., ... Ziegler, M. (2026, July 6). Artificial Intelligence and Individual Differences: A Perspective of the Differential Psychology, Personality Psychology, and Psychological Assessment Division of the German Psychological Society (DGPs). Retrieved from [osf.io/preprints/psyarxiv/j8szh\\_v2](https://osf.io/preprints/psyarxiv/j8szh_v2)
- Schoedel, R., & Mehl, M. (2024). Mobile sensing methods. In H. T. Reis, T. West, & C. M. Judd (Eds.), *Handbook of research methods in social and personality psychology* (3rd ed., pp. 297–321). Cambridge University Press. <https://doi.org/10.1017/9781009170123.014>
- SCHUHFRIED GmbH. (2022). Inventory for Testing Cognitive Abilities (INT). *Vienna Test System*.
- Stachl, C., Au, Q., Schoedel, R., Gosling, S. D., Harari, G. M., Buschek, D., ... & Bühner, M. (2020). Predicting personality from patterns of behavior collected with smartphones. *Proceedings of the National Academy of Sciences*, 117(30), 17680-17687. <https://doi.org/10.1073/pnas.1920484117>
- Stachl, C., Pargent, F., Hilbert, S., Harari, G. M., Schoedel, R., Vaid, S., ... & Bühner, M. (2020). Personality research and assessment in the era of machine learning. *European Journal of Personality*, 34(5), 613-631. <https://doi.org/10.1002/per.2257>
- Sust, L., Bergmann, M., Bühner, M., & Schoedel, R. (2026). Deep Beats, Deep Thoughts? Predicting General Cognitive Ability from Natural Music-Listening Behavior. *Journal of Intelligence*, 14(2), 29. <https://doi.org/10.3390/jintelligence14020029>
- Wrzus, C., & Neubauer, A. B. (2023). Ecologicalmomentary assessment: A meta-analysis on designs, samples, and compliance across research fields. *Assessment*, 30(3), 825–846. <https://doi.org/10.1177/10731911211067538>
- Youyou, W., Kosinski, M., & Stillwell, D. (2015). Computer-based personality judgments are more accurate than those made by humans. *Proceedings of the National Academy of Sciences*, 112(4), 1036-1040. <https://doi.org/10.1073/pnas.1418680112>