

Einsatz von Machine-Learning zur Vorhersage des Körperfettanteils

**Kann ein auf Forschungsdaten basiertes Machine-Learning-Modell eine
verlässliche Alternative zu den konventionellen Messverfahren des
Körperfettanteils darstellen?**

Eva Carvalho^{1,*}, Fabian Nadler¹, Shannon Schmuki¹

¹*Fachhochschule Graubünden*

**E-Mail Adressen: eva.carvalhonunes@stud.fhgr.ch,
fabian.nadler@stud.fhgr.ch,
shannon.schmuki@stud.fhgr.ch*

12. Dezember 2024

Inhaltsverzeichnis

Abstract.....	3
Einleitung.....	3
Stand der Forschung	3
Forschungsfrage und Methodik	3
Vorbereitung der Daten.....	4
Evaluierung potenzieller Modelle	4
Ermittlung der Top-Features	6
Trainieren der finalen Modelle	6
Testen der Modelle	7
Verknüpfung der Modellvorhersagen	8
Manuelle Optimierung	9
Resultate	10
Diskussion / Ausblick.....	10
Abbildungsverzeichnis.....	11
Tabellenverzeichnis	11
Literatur	11

1 Abstract

Diese erweiterte Version bietet eine ausführlichere und detaillierte Darstellung der Hauptarbeit und dient als Ergänzung zur kompakten Darstellung im Hauptdokument.

2 Einleitung

In dieser Arbeit steht die Entwicklung und Optimierung eines Machine-Learning-Modells zur Vorhersage des Körperfettanteils im Fokus. Dabei wurde ein Datensatz verwendet, der Messwerte wie Körperumfänge, Gewicht und Grösse von 252 erwachsenen Männern umfasst, ergänzt durch Referenzwerte aus hydrostatischem Wiegen. Ziel war es, ein präzises Modell zu erstellen, das eine kostengünstige und zugängliche Alternative zu traditionellen Methoden darstellt. Das Modell wurde durch die Auswahl geeigneter Algorithmen, wie lineare Regression und Random Forest Regressor, sowie durch die Optimierung der Hyperparameter schrittweise verbessert. Zusätzlich wurden die relevantesten Merkmale identifiziert, um die Modellkomplexität zu reduzieren und die Genauigkeit zu steigern. Abschliessend wurden die Vorhersagegenauigkeit und Robustheit des finalen Modells evaluiert. Traditionelle Methoden zur Messung des Körperfettanteils, wie DEXA-Scans oder Unterwasserwiegen, gelten als sehr präzise, sind jedoch aufwendig und teuer. Machine-Learning-Ansätze zeigen hier Potenzial, diese Herausforderungen zu überwinden. In dieser Arbeit wurde ein eigens entwickeltes Modell erstellt und trainiert, um eine präzise und praktikable Vorhersage des Körperfettanteils zu ermöglichen.

3 Stand der Forschung

Zur Analyse von Körperfett und Körperzusammensetzung gibt es bereits verschiedene beliebte Methoden. DEXA-Scans (Dual-Energy X-ray Absorptiometry) nutzen Röntgenstrahlen mit zwei Energieleveln, um präzise zwischen Fett-, Muskel- und Knochengewebe zu unterscheiden. Diese Methode wird häufig in der klinischen und sportmedizinischen Praxis eingesetzt (Geoffrey, 2024). Das Unterwasserwiegen bestimmt durch Messung des Gewichts an Land und unter Wasser die Körperdichte und darauf basierend den Fettanteil. Diese Methode gilt als Goldstandard in der Genauigkeit, erfordert jedoch aufwendige und kostenhohe Ausstattung (Wood, 2005). Eine weitere Option ist die Air Displacement Plethysmography (z. B. Bod Pod), die durch Luftverdrängung das Körpervolumen misst und so eine präzise Fettanteilsanalyse ermöglicht (Wood, 2004).

Neben diesen physikalischen Methoden existieren inzwischen bereits Machine-Learning-Modelle, die den Körperfettanteil vorhersagen können. Solche Modelle kombinieren oft Parameter wie Gewicht, Bauchumfang oder Hautfaltendicke, um mit Algorithmen wie Support Vector Regression (SVR) oder neuronalen Netzwerken genaue Vorhersagen zu treffen. Beispielsweise wurden hybride Ansätze wie SVR kombiniert mit Emotional Artificial Neural Networks entwickelt, die gegenüber klassischen Methoden eine vergleichbare oder sogar bessere Genauigkeit erreichen. Diese Modelle zeigen, dass moderne Technologien eine kosteneffiziente und praktikable Ergänzung zu traditionellen Messmethoden bieten können (Hussain et al., 2021).

4 Forschungsfrage und Methodik

Diese Arbeit untersucht, ob ein Machine-Learning-Modell auf Basis eines Forschungsdatensatzes in der Lage ist, den Körperfettanteil zuverlässig vorherzusagen. Ziel ist es, die Leistungsfähigkeit des Modells anhand gängiger Evaluationsmethoden zu testen und zu bewerten, ob es präzise Vorhersagen liefern kann.

Die momentan genauesten Methoden zur Bestimmung des Körperfettanteils wie DEXA-Scans, Unterwasserwiegen oder Air Displacement sind mit hohen Kosten und einem hohen logistischen Aufwand verbunden. Spezialisierte Geräte und geschultes Personal werden benötigt, was den Einsatz im Alltag oder in grossen Studien erschwert (Virgile, 2019). Zwar

existieren auch weniger aufwendige Methode, wie z.B. die Hautfaltenmessung, jedoch bieten diese eine geringere Genauigkeit (Fischl, 2021).

Ein Machine-Learning-Modell könnte eine Alternative bieten, die Komplexität sowie Kosten der Messungen reduziert und sie für eine breitere Zielgruppe zugänglich macht.

Diese Untersuchung beschränkt sich ausschliesslich auf Erwachsene Männer und berücksichtigt keine spezifischen gesundheitlichen Beeinträchtigungen, die die Ergebnisse verfälschen könnten.

4.1 Vorbereitung der Daten

Nach dem Import der Daten wurde nach einer ersten Analyse festgestellt, dass einige Masseinheiten angepasst werden sollten. Die Spalten „Weight“ und „Height“ wurden entsprechend ins metrische System umgerechnet. Anschliessend wurde überprüft, dass keine fehlenden Werte im Datensatz vorhanden sind bevor der Datensatz in Trainings- und Testdaten aufgeteilt wurde.

Um die Daten weiter zu untersuchen, wurden Scatterplots erstellt, um potenzielle Ausreisser zu identifizieren, welche die Güte des Modells beeinträchtigen könnten. Einige identifizierte Ausreisser wurden für das Training des Modells aus dem Datensatz entfernt. Zur Analyse der Korrelationen zwischen den Merkmalen wurde eine Heatmap erstellt. Dabei zeigten die Merkmale „Abdomen“, „Chest“, „Hip“ und „Weight“ die höchste Korrelation mit der Zielvariable. Um diese Korrelationen weiter zu veranschaulichen, wurden die relevantesten Merkmale zusätzlich in einer Scatter-Matrix dargestellt.

4.2 Evaluierung potenzieller Modelle

Im Anschluss wurde geprüft, welche Machine-Learning Modelle für diese Art Datensatz infrage kommen könnten. Offensichtlich handelt es sich um ein Regressionsproblem, da der Zielwert nicht ein Typ, sondern ein prozentualer Wert ist.

Nach der initialen Recherche für potenzielle Modelle, fiel die engere Auswahl auf folgende:

- **Linear Regression (LR)**

Dieses Modell ist eher einfach zu implementieren, sorgt aber bei linearen Zusammenhängen für eine effektive Vorhersage. Da in den zuvor betrachteten Scatterplots eine gewisse Linearität zwischen dem Zielwert und den am meisten korrelierenden Features festgestellt wurde, war dieses Modell eine vielversprechende Option

- **Entscheidungsbäume (Decision Tree Regressor)**

Bei der Betrachtung der Korrelationen wurde festgestellt, dass einige Features nur eine geringfügige Korrelierung zum Zielwert aufweisen. Da ein Entscheidungsbaum auch mit solchen Daten umgehen kann, wurde dieses Modell auch als Option aufgenommen.

- **Random Forest Regressor (RFR)**

Da das vorliegende Problem eher komplex ist und der Zielwert eine Kombination aus vielen verschiedenen Faktoren, war der Random Forest Regressor auch eine valide Option. Dieser soll sehr präzise Vorhersagen auch bei nichtlinearen Beziehungen vorweisen können. Die Möglichkeit zur nachträglichen Modifizierung der diversen Hyperparameter sprachen ebenfalls für dieses Modell

- **Gradient Boosting Modell (XGB Regressor)**

Dieses Modell ist ähnlich wie der Random Forest Regressor gut geeignet, für komplexere Vorhersagen und soll diesen mit seiner Präzision sogar übertreffen können. Deshalb wurde das Modell XGB Regressor ebenfalls in Betracht gezogen.

Diese vier gewählten Modelle weisen verschiedene Stärken und Schwächen auf, sollten jedoch grundsätzlich gut für den vorliegenden Datensatz geeignet sein.

Nachdem die Auswahl der Modelle also getroffen wurde, war es an der Zeit diese mit dem Trainingsdatensatz initial zu trainieren. Dafür wurden die Modelle auf alle Features trainiert. Um eine erste Analyse zur Präzision der diversen Modelle durchführen zu können, wurde für alle Modelle ein Kreuzvalidierungsverfahren angewandt. Dabei wurden besonders der RMSE und R2 Wert betrachtet. Die Resultate wurden in der untenstehenden Tabelle visualisiert.

Tabelle 1: Modelle vergleichen

Modell	RMSE	R2
Linear Regression	4.5461	0.6715
Random Forest Regressor	4.9147	0.6088
XGB Regressor	5.5656	0.5036
Decision Tree Regressor	7.2083	0.1218

Aus dieser initialen Auswertung konnte bereits gut festgestellt werden, welche Modelle sich besser oder schlechter für die Vorhersage dieses spezifischen Zielwertes eignen. Das Ziel dieses Vergleiches liegt darin, sich für die zwei am besten geeigneten Modelle zu entscheiden, damit diese zum Schluss kombiniert werden können.

Das Modell welches deutlich auffiel war der "Decision Tree Regressor". Dieses Modell lieferte einen R2 Wert von 12 Prozent ab, was so viel bedeutet wie dass es den Zielwert kaum bis hin zu gar nicht anhand der vorhandenen Features beschreiben konnte. Entsprechend wurde dieses Modell direkt aus der Auswahl entfernt.

Die anderen drei Modelle weisen allesamt ähnliche Werte auf, während die Lineare Regression am besten und der XGB Regressor am schlechtesten abschnitt. Da die Modelle XGB Regressor und Random Forest Regressor beide auf Entscheidungsbäumen basieren, fiel die Entscheidung auf das initial bessere Modell, also den Random Forest Regressor.

Da das Modell Linear Regression am besten abschnitt, wurde dieses ebenfalls als zu verwendendes Modell gewählt.

4.3 Ermittlung der Top-Features

Nun, da die Selektion der beiden besten Modelle abgeschlossen war, ging es darum diese noch weiter zu verbessern. Dafür wurden zuerst die besten Features der jeweiligen Modelle ermittelt. “Die besten” bedeutete in diesem Fall die Features, welche am meisten Einfluss auf den Vorhersagewert besitzen.

- Lineare Regression
 - Bei diesem Modell wurden die Features, welche den höchsten Koeffizienten besitzen selektioniert, da diese folglich den grössten Einfluss auf die Vorhersage haben.
- Random Forest Regressor
 - Um die besten Features dieses Modells zu ermitteln, wurde vorab ein GridSearch durchgeführt, damit bei der Optimierung mit den bestmöglichen Hyperparametern gearbeitet werden kann. Nachdem das bestmögliche RFR-Modell ermittelt wurde, konnten aus diesem die Gewichte der einzelnen Features extrahiert werden. Anhand dieser Liste wurden im Anschluss auch die besten Attribute ermittelt.

Nun das die wichtigsten Features pro Modell ermittelt wurden, konnte der Optimierungsprozess fortgesetzt werden.

4.4 Trainieren der finalen Modelle

Um die Komplexität des final kombinierten Modells zu reduzieren, wurde für das erneute Trainieren der Modelle eine Kombination, der jeweils vier stärksten gewichteten Features pro Modell verwendet. Dadurch konnten redundante Features eliminiert werden und beide Modelle verfügten trotzdem noch über ihre jeweiligen Top-Features.

Insgesamt führte dies zu einer geringeren Dimensionalität und die Gefahr von Overfitting konnte reduziert werden.

Die Modelle wurden auf folgende Features trainiert:

- Wrist, Weight, Neck, Chest, Height und Abdomen

Die neu trainierten Modelle wurden nun erneut auf die RMSE und R2 Werte geprüft, um festzustellen ob bereits eine gewisse Optimierung sichtbar ist. Der Vergleich ist in der untenstehenden Tabelle visualisiert.

Tabelle 2: Genauigkeit nach Training

Modell	RMSE vorher	RMSE nachher	R2 vorher	R2 nachher
Linear Regression	4.5461	4.3748	0.6715	0.6959
Random Forest Regressor	4.9147	4.8930	0.6088	0.6188

Wie in der Tabelle ersichtlich, ist bereits eine kleine Verbesserung im Vergleich zu vorher ersichtlich.

Da es sich bei der Linearen Regression nicht um ein sehr komplexes MLM handelt, konnte bei diesem Modell die Genauigkeit nicht mehr stark verbessert werden.

Deshalb rückte nun die Optimierung des RFR-Modells in den Vordergrund.

Eine wichtige Eigenschaft des Random Forest Regressor Modells sind die verschiedenen Hyperparameter.

Um die optimalen Parameter zu ermitteln konnte erneut das GridSearch Prinzip verwendet werden. In diesem Fall wurden nur mit den 2 wichtigsten Hyperparametern gearbeitet, da die restlichen abgesehen von einer verlängerten Laufzeit keine merkliche Verbesserung des Modells bewirkten.

Folgende zwei Hyperparameter wurden mit je fünf verschiedenen Optionen betrachtet:

Tabelle 3: Hyperparameter

n_estimators	max_features
100, 200, 300, 400, 500	3, 4, 5, 6, 7

Im Anschluss wurde das RFR-Modell mit den ermittelten optimalen Hyperparametern trainiert, welche folgende Werte betrugen:

- n_estimators: 100
- max_features: 5

Nach diesem Schritt waren die Modelloptimierungen abgeschlossen.

4.5 Testen der Modelle

Da die beiden Modelle Lineare Regression und Random Forest Regressor nun beide für diese Anwendung abschliessend optimiert wurden, war es an der Zeit diese auf die Vorhersagegenauigkeit zu überprüfen.

Dafür wurde der zu Beginn vom Rest abgespaltete Testdatensatz verwendet.

Von beiden Modellen wurde für alle Einträge des Testdatensatzes eine Vorhersage getroffen.

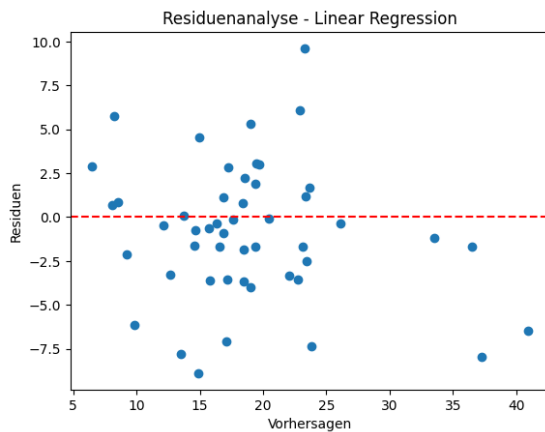
Um die Präzision der einzelnen Modelle zu prüfen, konnte nun erneut der RMSE und R2 Wert berechnet werden. Die Resultate dieses Tests sind in der folgenden Tabelle dargestellt.

Tabelle 4: Testresultate

Modell	RMSE	R2
Lineare Regression	3.9675	0.7176
Random Forest Regressor	3.7096	0.7531

Beide Modelle haben in dem Test verhältnismässig gut abgeschnitten.

Um zu prüfen, ob die Lineare Regression die Daten gut beschreibt und vertrauenswürdige Vorhersagen erstellt, wurde eine Residuen Analyse durchgeführt.



Wie in dieser Abbildung gut ersichtlich, konnte aus den getroffenen Vorhersagen kein systematisches Muster erkannt werden. Die Zielwerte schienen willkürlich und normal gestreut zu sein.

Aus dieser Erkenntnis konnte das Fazit gezogen werden, dass das Lineare Regression Modell Qualitative Annahmen trifft und die Daten gut beschreibt. Entsprechend wurde dieses Modell weiterhin verwendet.

4.6 Verknüpfung der Modellvorhersagen

Zu Beginn wurde das Ziel definiert, die zwei optimierten Modelle zu verknüpfen. Es wurden verschiedene Verfahren zur Realisierung betrachtet. Die Entscheidung fiel auf das “Weighed Averaging” Verfahren, dieses ist sehr simpel und doch wirksam.

Mit Hilfe einer Optimierungsfunktion wurden die optimalen Gewichte für die jeweiligen Modelle ermittelt.

Diese betrugen folgende Werte:

Tabelle 5: optimale Gewichte

Modell	Gewicht
Lineare Regression	22.32%
Random Forest Regressor	77.67%

Die daraus entstandenen kleinsten Gütewerte betrugen:

RMSE: 3.6855

R2: 0.7563

Im Vergleich zu den vorherigen Ergebnissen konnte durch die Kombination der beiden Modelle das Finale Ergebnis zwar nur minimal optimiert werden, das Modell sollte aber einiges stabilere Vorhersagen treffen können, da nun ein Lineares sowie ein Nicht-Lineares Modell in Betracht gezogen wird.

4.7 Manuelle Optimierung

Durch die Kombination der Modelle bestand die Möglichkeit, dass die Hyperparameter des RFR-Modells nicht mehr optimal waren, da es nicht mehr allein für die Vorhersagen verantwortlich war. Daher wurden verschiedene Hyperparameter-Variationen getestet, um das Modell weiter zu optimieren.

Tabelle 6: Hyperparameter vergleichen

N_estimators	Max_features	RMSE (Kombiniert)	R2 (Kombiniert)
100	6	3.6652	0.7589
200	6	3.6790	0.7571
300	6	3.6933	0.7552
100	5	3.6855	0.7563
200	5	3.6332	0.7631
300	5	3.6583	0.7598
100	4	3.5898	0.7688
200	4	3.5801	0.7700
300	4	3.5982	0.7677
100	3	3.6757	0.7575
200	3	3.6476	0.7612
300	3	3.6351	0.7629

Mit Hilfe von diesen Tests konnte das kombinierte Modell noch weiter optimiert werden. Wie in der Tabelle ersichtlich, war der RMSE-Wert sowie der R2 Werte bei den folgenden Hyperparametern am idealsten:

n_estimators: 200

max_features: 4

Im direkten Vergleich mit den vorherigen Hyperparametern ist die Optimierung klar ersichtlich:

Tabelle 7: Übersicht Optimierung

	Vorher	Nachher	Differenz
RMSE	3.6855	3.5801	- 0.1054
R2	0.7563	0.7700	+ 0.0137

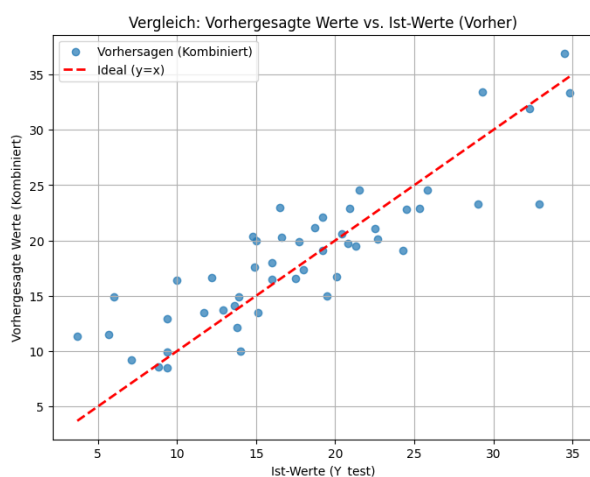


Abbildung 1: Qualitätsvergleich vorher

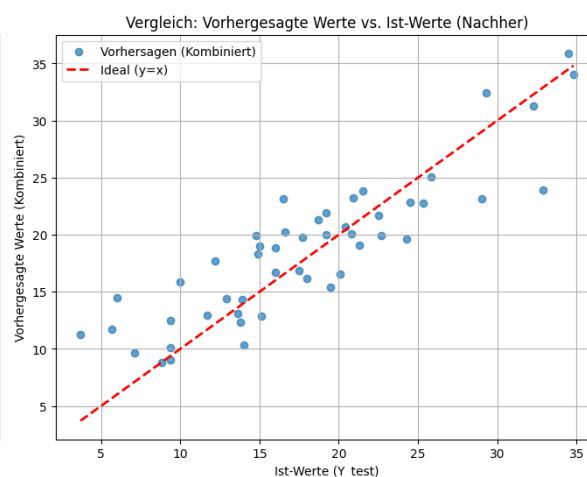


Abbildung 2: Qualitätsvergleich nachher

5 Resultate

Das finale Modell ist in der Lage, den Körperfettanteil von erwachsenen Männern anhand von Messwerten am Körper vorherzusagen. Als Eingabevariablen dienen der Umfang des Unterleibs, der Brust, des Halses und des Handgelenks sowie Körpergrösse und Gewicht. Die Bewertung der Güte des Modells ergab eine durchschnittliche Abweichung (RMSE) von ca. 3.58% und einen Determinationskoeffizienten (R^2) von 0,77. Diese Ergebnisse zeigen, dass das Modell in der Lage ist, einen Grossteil der Varianz im Körperfettanteil zu erklären, wobei dennoch ein gewisser Anteil an Vorhersageungenauigkeit bleibt. Die Kombination der beiden Modellansätze, des linearen Regressionsmodells und eines Random Forests, hat zu einer robusten Vorhersageleistung beigetragen, indem die Stärken beider Methoden genutzt wurden.

6 Diskussion/ Ausblick

Die Ergebnisse des Modells zeigen eine solide Leistung bei der Vorhersage des Körperfettanteils. Es bleibt jedoch ein Anteil an Vorhersagefehlern, was auf Modellschwächen oder mögliche Einflüsse nicht berücksichtigter Faktoren hinweisen kann.

Ein möglicher Grund für die verbleibende Ungenauigkeit könnten nicht berücksichtigte Faktoren wie Lebensstil, genetische Faktoren oder die Körperzusammensetzung sein.

Die Kombination aus linearem Regressionsmodell und Random Forest hat sich als vielversprechend erwiesen, da sie sowohl lineare als auch nichtlineare Beziehungen in den Daten berücksichtigt.

Dennoch könnten alternative Modelle wie neuronale Netze oder gradientenbasierte Boosting-Ansätze potenziell noch bessere Ergebnisse erzielen, insbesondere bei grossen und vielfältigeren Datensätzen.

Praktisch könnte das Modell eine kostengünstige und leicht zugängliche Alternative zu bestehenden Messmethoden bieten, insbesondere wenn teure und spezialisierte Geräte nicht zur Verfügung stehen. Wird das Modell jedoch ausserhalb der Trainingspopulation eingesetzt, ist Vorsicht geboten, und es bedarf einer weiteren Validierung an unabhängigen Datensätzen, um die Robustheit und Übertragbarkeit zu prüfen.

Zukünftige Arbeiten könnten sich zusätzlich auf eine Erweiterung des Datensatzes auf eine vielfältigere Zielgruppe, die Nutzung weiterer Merkmale sowie die Anwendung weiterer Modelle konzentrieren. Dadurch könnte das Modell generalisiert und auf eine grössere Zielgruppe ausgelegt werden.

Abbildungsverzeichnis

Abbildung 1: Qualitätsvergleich vorher	Abbildung 2: Qualitätsvergleich nachher.....	9
--	--	---

Tabellenverzeichnis

Tabelle 1: Modelle vergleichen	5
Tabelle 2: Genauigkeit nach Training	6
Tabelle 3: Hyperparameter	7
Tabelle 4: Testresultate	7
Tabelle 5: optimale Gewichte	8
Tabelle 6: Hyperparameter vergleichen.....	9
Tabelle 7: Übersicht Optimierung	9

Literatur

Capol, C. (2022). *Cds-templates/term-paper* [GitHub]. GitHub.

<https://github.com/muncca/cds-templates/tree/main/term-paper>

ChatGPT. (o. J.). Korrektur der Grammatik. Abgerufen 10. Dezember 2024, von

<https://chatgpt.com>

Fischl, M. (2021, Februar 22). *KAF Akadamie*. Körperfettanteil messen: 5 Arten und deren Vor- & Nachteile | B Lizenz Thema. <https://kaf-akademie.de/>

Geoffrey, R. (2024, März 11). *Bone Densitometry (DEXA , DXA)*. Radiologyinfo.Org.

<https://www.radiologyinfo.org/en/info/dexa>

Géron, A., Rother, K., & Demmig, T. (2023). *Praxiseinstieg Machine Learning Mit Scikit-Learn, Keras und TensorFlow: Konzepte, Tools und Techniken Für Intelligente Systeme*.

O'Reilly Verlag GmbH & Co. KG. <http://ebookcentral.proquest.com/lib/fh-ostschweiz/detail.action?docID=7285794>

Hussain, S. A., Cavus, N., & Sekeroglu, B. (2021). Hybrid Machine Learning Model for Body Fat Percentage Prediction Based on Support Vector Regression and Emotional Artificial Neural Networks. *Applied Sciences*, 11(21), Article 21.

<https://doi.org/10.3390/app11219797>

Thompson, D. (2024, April 9). *The Science Behind DEXA Scans: How & Why It Works*.

Meliora Integrative Medicine. <https://melioramed.com/articles/the-science-behind-dexa-scans-how-it-works-and-why-it-works/>

Virgile, A. (2019, Januar 31). Body Composition in Sport: BOD POD and Underwater Weighing. *Adam Virgile Sports Science*. <https://adamvirgile.com/2019/01/31/body-composition-in-sport-densitometric-methods/>

Wood, R. (2004, Februar). *Whole-Body Air-Displacement Plethysmography*. Whole-Body

Air-Displacement Plethysmography. <https://www.topendsports.com/testing/tests/bodpod.htm>

Wood, R. (2005, April). *Hydrostatic Underwater Weighing*. Hydrostatic Weighing for

Measuring Body Composition. <https://www.topendsports.com/testing/tests/underwater.htm>

Anhang

GitHub: ausführliche Version, Jupyter Notebook und Datensätze:
https://github.com/nadlerfabian/CDS_1011_LN2