

Machbarkeitsstudie

Modellbildung zur Bestimmung von Körperhaltungen anhand von Belastungen eines Lattenrosts

erstellt von

Karolina Ochs & Daniel Mansfeldt

Inhaltsverzeichnis

1	Einleitung	1
2	Ausgangslage	1
2.1	Rohdaten	1
2.2	Zusammenhang zwischen Messwerten und Liegepositionen	2
2.3	Relevante Körperpunkte	2
3	Datenaufbereitung	3
4	Bewertungsmetriken	4
5	Modellbildung	4
5.1	Vorauswahl von Methoden	4
5.2	Modelle	5
5.2.1	Support Vector Machine	5
5.2.2	XGBoost	5
5.2.3	Multilayer Perceptron	6
6	Ergebnisse und Diskussion	7
7	Fazit	9
	Literatur	11

1 Einleitung

Das Unternehmen *Laromed* aus Schleswig, Schleswig-Holstein, entwickelt und vertreibt Systeme, die in der Altenpflege zum Einsatz kommen. Dazu gehören Lattenroste, Matratzen und vollständige Pflegebetten. Fokus bei all diesen Produkten liegt auf besonderen Funktionalitäten, die sowohl die Lebensqualität der Patienten erhöhen als auch das Pflegepersonal bei ihrer Arbeit unterstützen.

Die Liegehaltung ist – insbesondere in der Pflege – ein wichtiges Merkmal von Patienten. Mit Kenntnis über sie lassen sich verschiedene Abläufe effizienter und für die Patienten angenehmer gestalten. *Laromed* hat deswegen ein Interesse daran, die Liegehaltung von Patienten kontinuierlich und zuverlässig bestimmen zu können.

Zu diesem Zweck hat Laromed ein sensorisches Lattenrost entwickelt, dessen Latten auf je zwei Kraftsensoren gelagert sind. Diese erlauben es verschiedene Belastungen zu messen, die über eine Matratze auf das Lattenrost wirken. Je nach Liegeposition ergeben sich unterschiedliche Belastungen, sodass ein Rückschluss auf die Körperhaltung der darauf liegenden Person möglich wird. Allerdings ist ohne eine genauere Betrachtung unklar wie korrekt diese Abbildung von Kraftwerten zu Liegepositionen umgesetzt werden kann.

Somit ist das Ziel dieser Arbeit die Betrachtung verschiedener Methoden zur Erzeugung von Modellen, welche die beschriebene Abbildung vornehmen. Dabei gilt es zu prüfen, ob eins oder mehrere Modelle geeignet sind die Liegepositionen zuverlässig und ausreichend genau vorherzusagen. Um hierüber eine Aussage zu treffen und verschiedene Modelle miteinander zu vergleichen, ist ebenfalls eine Metrik zur Bewertung der Modelle zu definieren.

2 Ausgangslage

2.1 Rohdaten

Grundlage dieser Arbeit ist ein Datensatz der von der Firma *Laromed* erstellt worden ist.

Der Datensatz setzt sich aus zwei Teilen zusammen. Der erste Teil enthält die Koordinaten der einzelnen Körperpunkte (hier auch Joints genannt). Aufgezeichnet wurden diese mit Hilfe eines *Azure Kinect* Systems der Firma *Microsoft*. Das System definiert 32 Körperpunkte, die jeweils in Form eines eigenen Koordinatensystems aufgefasst werden. Position und Orientierung der Körperpunkte werden relativ zu dem globalen Koordinatensystem der *Azure Kinect* bestimmt. Die Kinect-Kamera – und damit das globale Koordinatensystem – ist mittig und senkrecht auf die Liegeebene gerichtet.

Der zweite Teil enthält die Messwerte der Kraftsensoren, auf denen die Latten des Lattenrostes gelagert sind. Für jede Beobachtung liegen 64 Kraftwerte vor. Die Werte liegen umgerechnet in der Einheit Gramm vor.

2.2 Zusammenhang zwischen Messwerten und Liegepositionen

Für die Bestimmung liegender Körperhaltungen reicht es aus, die Lage der Körperpunkte in der Ebene und ohne Rotation zu betrachten. Dies hat folgende Gründe:

1. Die *Orientierung* eines Körperpunktes hat entweder keinen bedeutsamen Einfluss auf die Liegehaltung oder falls doch, dann wird dies ausreichend durch die kartesischen Koordinaten beschrieben. Die Drehung eines Handgelenks beispielsweise verändert die Körperhaltung in irrelevanter Weise für die Liegehaltung. Ändert sich hingegen die Orientierung eines Hüftgelenks, dann hat dies eine große Auswirkung auf die Körperhaltung insgesamt. Allerdings wird diese Änderung in gleichem Ausmaß durch die Position der Körperpunkte beschrieben, die mit dem Hüftgelenk eine kinematische Kette bilden.

Außerdem kann auch die Orientierung einer Seitenlage allein durch die Position der Körperpunkte in der Ebene bestimmt werden, weil deren Position relativ zueinander charakteristisch für eine gewisse Orientierung ist. Befinden sich beispielsweise die Handgelenke auf der linken Seite (Draufsicht, Kopf oben) ist dies ein starkes Indiz für eine Lage auf der rechten Körperseite.

2. Die Ausrichtung der Kinect-Kamera erlaubt es die Höhenkoordinate (z-Achse) zu vernachlässigen. Für eine liegende Person lassen sich nämlich alle Körperpunkte näherungsweise in einer Ebene annehmen. Diese Ebene ist parallel zur Oberfläche der Matratze; zu welcher die z-Achse des globalen Koordinatensystems wiederum orthogonal ausgerichtet ist. Unter diesen Annahmen enthält die z-Koordinate keine Informationen über die Position der Körperpunkte zueinander. Diese Annäherung ist selbst dann anwendbar, wenn die Person in Seitenlage liegt. Grund dafür ist, dass Unterschiede in der Höhenkoordinate sich auch in den Positionen in der x-y-Ebene widerspiegeln. So liegen bei Rückenlage die Schulterpunkte in der Ebene weiter auseinander als in Seitenlage. Die Distanz zwischen den Schulterpunkten kann so schon als Indikator für eine Bauch- bzw. Rücklage oder eine Seitenlage dienen.

2.3 Relevante Körperpunkte

Das Azure Kinect System definiert 32 Körperpunkte, wie in Abb. 1 zu sehen ist. Um die Liegehaltung einer Person ausreichend gut zu bestimmen, sind nicht alle davon bedeutsam. Deswegen können ausgewählte Körperpunkte aus dem Datensatz entfernt werden und so die Problemstellung vereinfacht werden.

In Abb. 1 sind alle Körperpunkte, die für die weitere Betrachtung vernachlässigt werden, rot markiert. Die Positionen all dieser Körperpunkte werden maßgeblich durch die restlichen Körperpunkte beeinflusst und tragen demnach nicht bedeutsam zur Beschreibung der Liegehaltung bei.

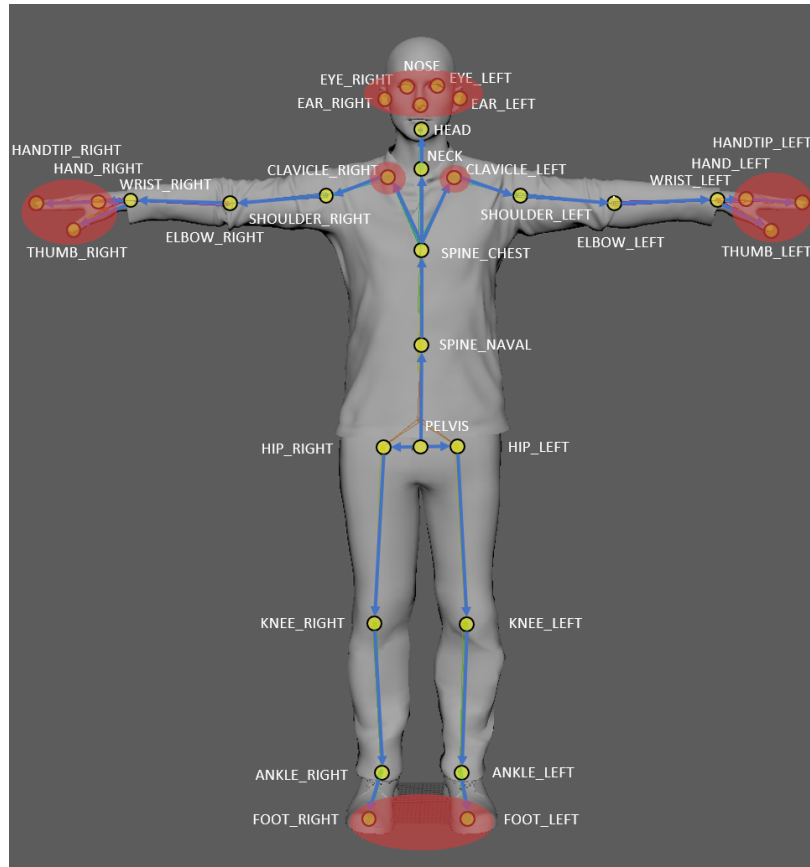


Abbildung 1: Körperpunkte nach Definition des Azure Kinect Systems. Die rot markierten Körperpunkte sind irrelevant für die Bestimmung von Liegehaltungen. Quelle: <https://docs.microsoft.com/en-us/azure/kinect-dk/body-joints>

3 Datenaufbereitung

Um den Datensatz für das Training der Modelle vorzubereiten werden zunächst sämtliche, zu vernachlässigende Merkmale entfernt. Eingabe- und Zielwerte werden anschließend separat auf einen Wertebereich von -1 bis 1 normalisiert. Trotz der einheitlichen Einheiten der Merkmale ist dies sinnvoll, um das Training der Modelle zu stabilisieren.

Die Aufteilung des Datensatzes in Trainings-, Test- und ggf. Validierungsdatsatz erfolgt individuell für die jeweilige Methode.

4 Bewertungsmetriken

Um die Modelle bewerten und vergleichen zu können, ist es notwendig eine geeignete Metrik zu definieren. Bei der vorliegenden Problemstellung ist letztendlich entscheidend ob eine gewisse Körperhaltung *als ganzes* ausreichend gut approximiert werden konnte. Dies ist der Fall, wenn die Positionen aller Körperpunkte einzeln ausreichend genau bestimmt werden konnten. Moderate Abweichungen zwischen Vorhersagewert und Zielwert der einzelnen Körperpunktpositionen verzerren die Vorhersage einer Körperhaltung nur unwesentlich. Weicht jedoch mindestens ein Körperpunkt zu stark von seinem Zielwert ab, ändert dies die Körperhaltung bedeutsam. Für ein Fußgelenk beispielsweise kann dies den Unterschied zwischen einem gestreckten oder angewinkeltem Bein ausmachen; moderate Abweichungen tun dies nicht.

Aus den genannten Überlegungen wird eine Metrik abgeleitet, die eine Aussage darüber erlaubt welcher Anteil an betrachteten Körperhaltungen *korrekt* aus den Eingabewerten gefolgert wurde.

Die Position eines Körperpunktes gilt als ausreichend gut approximiert, wenn die Distanz zwischen ihm und dem zugehörigen Zielpunkt kleiner als 50 mm ist. Eine Körperhaltung gilt dann als korrekt bestimmt, wenn alle zugehörigen Körperpunkte in ihren jeweiligen Toleranzbereich fallen.

Die Größe der Toleranz von 50 mm ist frei gewählt. Dennoch berücksichtigt sie folgende Annahmen, die einen Einfluss auf die Qualität einer Vorhersage haben.

- Es wird angenommen, dass sich in Bauch- oder Rückenlage benachbarte Körperpunkte, selbst im Grenzfall nicht überdecken können.
- Es wird angenommen, dass Abweichungen der Körperpunktpositionen im Toleranzbereich, die Körperhaltung nicht bedeutsam ändern.
- Der Toleranzbereich ist groß genug, damit eine möglichst hohe Genauigkeit bei der Bestimmung der Körperhaltungen erreicht werden kann, aber dennoch aussagekräftige Vorhersagen zulässt.

Distanzen zwischen Zielposition und vermuteter Position, die kleiner als 50 mm sind werden fortan als *moderat* bezeichnet.

5 Modellbildung

5.1 Vorauswahl von Methoden

Prinzipiell sind alle Methoden anwendbar, die eine Regression ermöglichen. Im Rahmen dieser Arbeit kann allerdings nur eine Teilmenge der anwendbaren Methoden untersucht werden. Somit wurden folgende drei Methoden ausgewählt, die sich in der Literatur als erfolgversprechend herausgestellt haben:

- Support Vector Machines (SVM) [2]
- eXtreme Gradient Boosted Trees (XGBoost) [1]
- Multilayer Perceptron (MLP) [3]

Diese Methoden werden verwendet, um verschiedene konkrete Modelle zu erstellen, die in der Lage sind für gegebene Belastungen die zugehörige Liegeposition vorherzusagen. Die Möglichkeit verschiedene Modelle mit einer Methode zu erzeugen rührt daher, dass die konkrete Ausprägung einer Methode mit sogenannten Hyperparametern gesteuert werden kann. Außerdem besitzen einige Algorithmen auch zufällige Abläufe sodass selbst mit ein und der selben Parametrisierung verschiedene Modelle entstehen können. Sofern eine Methode verwendet wurde, um mehrere Modelle zu erzeugen, wird schlussendlich mit Hilfe der in Abschnitt 4 vorgestellten Metrik ein finales Modell ausgewählt, welches die besten Ergebnisse liefert. Dabei wird ein separater Teil des Datensatzes, der sogenannte Validierungsdatsatz verwendet, der nicht zur Erzeugung der Modelle genutzt wurde. Anschließend können die finalen Modelle der drei Methoden gegenüber gestellt werden, um das am besten geeignete Modell für die vorliegende Aufgabenstellung auszuwählen.

5.2 Modelle

5.2.1 Support Vector Machine

Die Idee hinter der Regression mit einer Support Vector Machine ist der Einsatz einer sogenannten *Support Vector Regression* der Eingabedaten nach der Transformation in einen höherdimensionalen Raum. Eine *Support Vector Regression* funktioniert ganz ähnlich wie eine lineare Regression, mit dem Unterschied, dass nicht alle Punkte den Verlauf der Geraden bestimmen, sondern nur eine Teilmenge, die sogenannten *Support Vectors*. Bei der vorliegenden Problemstellung handelt es sich aufgrund der 64 Eingabewerte sogar um eine *Multivariate Support Vector Regression*, sodass keine Gerade für die Regression gesucht wird, sondern eine 64-dimensionale Hyperebene. Durch die erwähnte Transformation in einen höherdimensionalen Raum wird die Regression durch die Hyperebene überhaupt erst möglich.

Um eine SVM für die vorliegende Aufgabenstellung zu erzeugen, also 34 Jointkoordinaten vorherzusagen, müssen 34 separate SVMs erzeugt werden, da diese je nur die Vorhersage von einem Ausgabewert zulassen. Bei dieser Methode wurde keine weitere Optimierung der Hyperparameter vorgenommen, die Parameter wurden auf ihren Defaultwerten belassen und können der Dokumentation von scikit-learn entnommen werden.

5.2.2 XGBoost

Bei XGBoost handelt es sich um einen effizienten Algorithmus zur Erzeugung eines Ensembles von Entscheidungsbäumen, sogenannten *Gradient Boosted Trees*. Hierbei

meint *Boosting* eine spezielle Technik zur Erzeugung und Anwendung der Entscheidungsbäume, die vorschreibt, dass die Bäume sequentiell in einer festgelegten Reihenfolge erzeugt und angewendet werden. Als Ergebnis kann jeder Baum die Schwächen des vorangegangenen Baums ausgleichen und die Ausgabe weiter verbessern. Beim *Gradient Boosting* werden die konkreten Parameter der einzelnen Bäumen basierend auf dem Gradienten einer *Loss-Funktion* festgelegt. Die Loss-Funktion beschreibt hierbei die aktuelle Güte des Ensembles, meist mit Hilfe der Differenz aus vorhergesagten und wahren Daten. Mit Hilfe des Gradienten ist es möglich diese Funktion zu optimieren, also ein Entscheidungsbaum-Ensemble zu erzeugen, welches die geringst mögliche Abweichung von den wahren Daten aufweist.

Wie bereits im Abschnitt 5.2.1 beschrieben, müssen auch für XGBoost 34 separate Entscheidungsbaum-Ensembles erzeugt werden. Für XGBoost wurde ebenfalls keine weitere Optimierung der Hyperparameter vorgenommen und die Parameter können der Dokumentation von scikit-learn entnommen werden.

5.2.3 Multilayer Perceptron

Ein MLP wird auch als vorwärtsbetriebenes neuronales Netz (engl. Feedforward Neural Network) bezeichnet und besteht aus einer Menge von künstlichen Neuronen, die in aufeinanderfolgenden Schichten angeordnet sind. Dabei wird die erste Schicht als *Eingabeschicht* und die letzte Schicht als *Ausgabeschicht* bezeichnet.

Alle Schichten dazwischen werden auch *versteckte Schichten* genannt. Die verschiedenen Schichten verarbeiten die eingegebenen Informationen (Belastungen), sodass die korrekten Ausgabedaten (Liegepositionen) erzeugt werden können. Hierzu besitzt jede Schicht eine Menge von Parametern, die zunächst in einem iterativen Lernprozess sinnvoll festgelegt werden müssen. Die verschiedenen möglichen Ausprägungen eines MLPs, werden wie in Abschnitt 5.1 beschrieben durch Hyperparameter wie die Anzahl Schichten, die Anzahl Neuronen oder weitere Charakteristika der Neuronen festgelegt. Ein konkretes Modell entsteht dann, wenn für eine festgelegte Ausprägung ein Lernprozess angestoßen wird, sodass zur Ausprägung auch eine Menge angelernter Parameter existiert. Der Lernprozess besitzt ebenfalls eine Menge verschiedener Hyperparameter. Da innerhalb des Lernprozesses mehrere zufällig gesteuerte Abläufe enthalten sind, können für ein und die selbe Ausprägung und identische Hyperparameter des Trainingsprozesses verschiedene Modelle erzeugt werden.

Um eine geeignete Konstellation von Hyperparametern für die vorliegende Problemstellung festzulegen, wird eine zufällige Suche auf einem empirisch definierten Hyperparameterraum durchgeführt; ein sogenanntes Hyperparametertuning. Die Anzahl der

Hyperparameter	Wert
Anzahl versteckte Schichten	4
Neuronen pro Schicht	60
Trainingsepochen	2000
Batchgröße	10
Lernrate	0,3

Tabelle 1: Set von Hyperparametern für die das MLP-Modell die größte Genauigkeit erreicht

Eingabeneuronen ist entsprechend der Anzahl der Kraftwerte auf 64 fixiert. Äquivalent beträgt die Anzahl der Neuronen in der Ausgabeschicht 34. Die im Rahmen des Hyperparametertunings angelernten Modelle werden jeweils bezüglich ihrer erreichten Genauigkeit (s. Abschnitt 4) auf dem Validierungsdatensatz verglichen. Als beste Kombination der Hyperparameter wird diejenige definiert für die das Modell die größte Genauigkeit erreicht. Diese Hyperparameter können Tabelle 1 entnommen werden.

6 Ergebnisse und Diskussion

Tabelle 2 zeigt für das beste Modell der jeweiligen Methode, die erreichten Werte der Genauigkeit A und die durchschnittliche Distanz $\bar{\delta}$ zwischen Vorhersage und Zielposition der Körperpunkte. Das MLP-Modell performt über alle Metriken hinweg am besten.

Model	A in %	$\bar{\delta}$ in mm
SVM	8,60	37,99
XG-Boost	37,26	25,48
MLP	61,70	20,32

Tabelle 2: Performannewerte der ML-Modelle

Um außerdem einen visuellen Eindruck über die Qualität der Vorhersagen der Modelle zu erhalten, sind in Abbildung 2 jeweils acht Vorhersagen pro Modell zu finden. Hier ist zu erkennen, dass die Vorhersage der Liegeposition insgesamt gut funktioniert und nur geringe Abweichungen von den wahren Werten vorliegen. Dabei ist insbesondere die Vorhersage der Position der Hände häufiger inkorrekt. Das hängt wahrscheinlich damit zusammen, dass das geringe Gewicht der Hände nur einen geringen Einfluss auf die Belastung der Kraftsensoren hat. Darüber hinaus ist auch auf einzelnen Bildern zu erkennen, dass die Target-Daten nicht ganz korrekt sind. Auf dem zweiten Bild von rechts, jeweils in der unteren Zeile beispielsweise, entspricht die gegebene Position des rechten Armes nicht dem Foto und die Vorhersagen der Modelle, wären eigentlich korrekt. Um die Güte der Modelle korrekt zu beurteilen, sollte diesbezüglich die Qualität der Daten verbessert werden.

Darüber hinaus scheint die erreichte Güte der Modelle optisch beurteilt höher, als die formal errechnete Genauigkeit von 61,7%. Diese Beobachtung wirft die Frage auf, ob die zuvor eingeführte Metrik ein geeignetes Maß für die Bewertung der Güte der Modelle darstellt. Diese Frage kann nur in Abstimmung mit Laromed beantwortet werden. Grundsätzlich wäre jedoch eine Erweiterung des Toleranzbereiches, eine Definition spezifischer Toleranzbereiche pro Joint oder das Zulassen von mehr als einem nicht akkuraten Joint denkbar.

Final gilt es zu klären, ob die erreichte Güte des MLP-Modells für die praktische Anwendung ausreicht. Diese Frage kann im Rahmen der Machbarkeitsstudie nicht

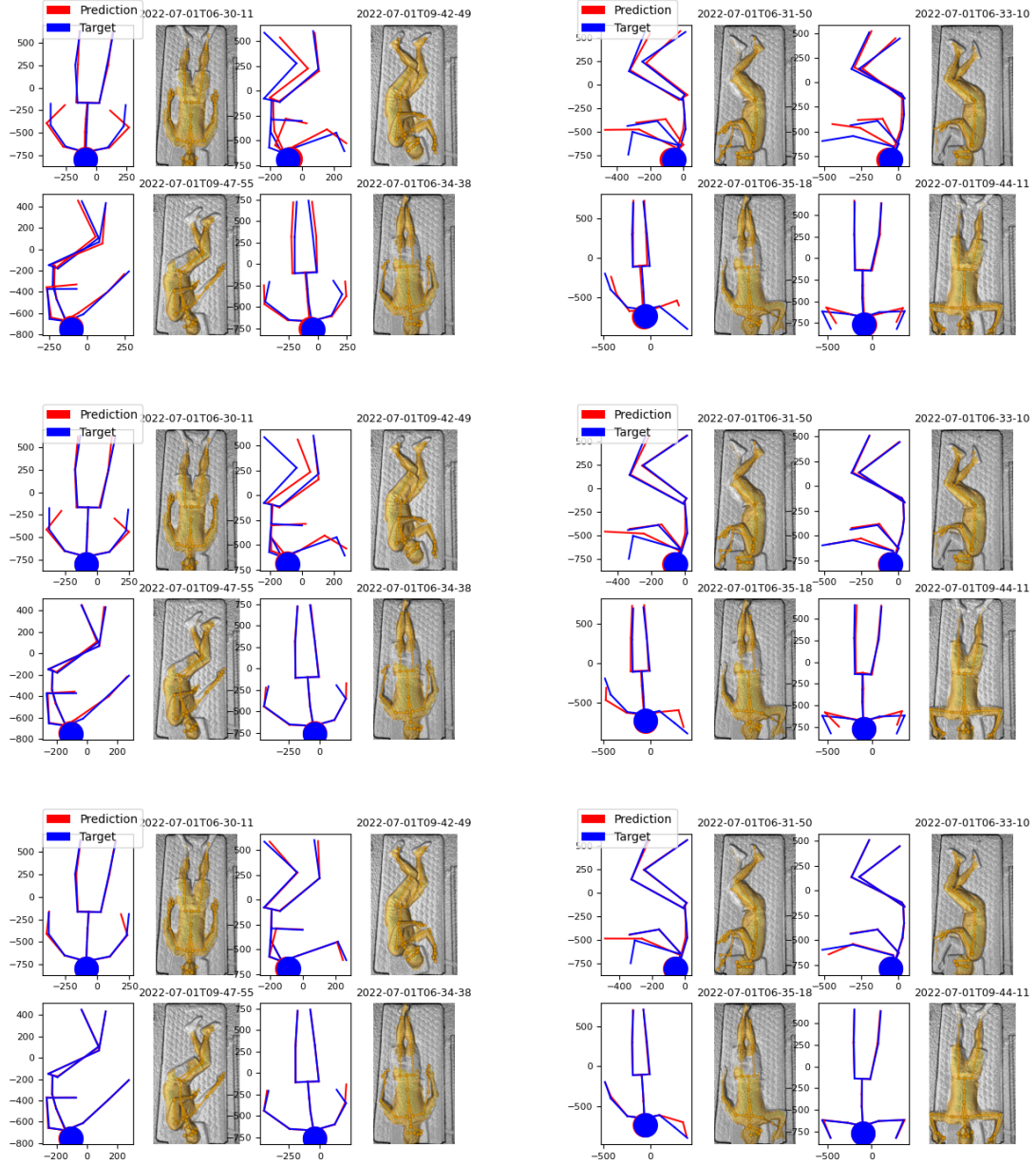


Abbildung 2: Vergleich zwischen den Target Daten und den Vorhersagen der Modelle.
Oben: SVM. **Mitte:** XGBoost. **Unten:** MLP.

abschließend geklärt werden, da hierzu Experten von Laromed basierend auf den Anforderungen des geplanten Gesamtsystems eine finale Einschätzung treffen müssen. Falls eine höhere Genauigkeit erfordert wird, bieten sich folgende Ansatzpunkte an:

1. Größerer Datensatz

Der verwendete Datensatz ist mit 1569 Beobachtungen für Methoden des maschinellen Lernens vergleichsweise klein. Es wird daher vermutet, dass sich die Treffsicherheit der Modelle mit einem größeren Datensatz weiter verbessern lässt. Dieser sollte für ein möglichst breites Spektrum an Körpertypen möglichst viele gängige Körperhaltung enthalten, um sicherzustellen, dass in der praktischen Anwendung für einen beliebigen Patienten eine möglichst akkurate Bestimmung der Körperhaltung möglich ist.

2. Höhere Qualität des Datensatzes

Neben der nicht vollständig korrekten Bewertung der Güte der Modelle durch einzelne inkorrekte Datenpunkte, wirkt sich dieser Umstand auch auf die Güte selbst aus. Das liegt darin begründet, dass die inkorrekten Daten auch für das Training der Modelle verwendet werden und diese somit lernen inkorrekte Vorhersagen zu treffen. Eine Erhöhung der Qualität der Daten hat somit auch direkt eine Steigerung der Qualität der Modelle zur Folge.

3. Hyperparametertuning

Wie bereits in Abschnitt 5.2.1 und Abschnitt 5.2.2 erwähnt, wurde für die Erzeugung der SVM- und XGBoost-Modelle kein Tuning der Hyperparameter durchgeführt. Die Implementierung eines Hyperparametertunings birgt somit noch Potenzial zur Verbesserung der erzeugten Modelle. Gleiches gilt für das Multilayer Perceptron, auch wenn bereits ein Hyperparametertuning durchgeführt wurde. Hier birgt insbesondere die Erweiterung des Suchraums (bspw. um weitere Aktivierungsfunktionen oder Loss-Funktionen) Potenzial für die Erzeugung korrekterer Modelle.

7 Fazit

Im Rahmen dieser Machbarkeitsstudie konnten mehrere Modelle erzeugt werden, die in der Lage sind basierend auf Auflagebelastungen des Lattenrosts die Körperposition einer darauf liegenden Person vorherzusagen. Zur Bewertung der Genauigkeit dieser Vorhersagen und zum Vergleich der Modelle konnte außerdem eine Metrik entwickelt werden. Gemäß dieser, ist von den untersuchten Methoden das Multilayer Perceptron am besten geeignet und erreicht eine Genauigkeit von maximal 62%. Auch wenn diese Genauigkeit auf den ersten Blick nur moderat wirkt, zeigt sich bei visueller Darstellung der Vorhersagen, dass diese sehr dicht an die korrekten Liegepositionen herankommen. Somit bleibt die finale Einschätzung über die Güte der Modelle und die Eignung für den praktischen Einsatz bei der Firma Laromed. Gegebenenfalls kann

die Metrik zukünftig noch adaptiert werden, sodass die empfundene Güte der Modelle besser in eine zugehörige Maßzahl übersetzt werden kann.

Perspektivisch kann die Qualität der Modelle beispielsweise durch eine Vergrößerung des Datensatzes oder eine Erhöhung der Qualität des Datensatzes erzielt werden. Außerdem kann auch eine Erweiterung des Hyperparametertunings die Güte der Modelle erhöhen.

Literatur

- [1] T. Chen und C. Guestrin. „Xgboost: A scalable tree boosting system“. In: *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*. 2016, S. 785–794.
- [2] C. Cortes und V. Vapnik. „Support-vector networks“. In: *Machine learning* 20.3 (1995), S. 273–297.
- [3] M. Minsky und S. A. Papert. *Perceptrons, Reissue of the 1988 Expanded Edition with a new foreword by Léon Bottou: An Introduction to Computational Geometry*. MIT press, 2017.