

## Eine vergleichende Studie von Modellen des Maschinellen Lernens

# Prädiktionsgüte der Rückfederung in der Blechbearbeitung

P. Lange, P. Kurrle, U. Pado

**ZUSAMMENFASSUNG** Die Rückfederung ist ein für die Fertigungsgenauigkeit maßgeblicher Parameter bei Biegeprozessen. In dieser Studie wurde die Eignung von maschinellen Lernmethoden zur Vorhersage der Rückfederung geprüft und die Prädiktionsgüte der Vorhersagen evaluiert. Für den betrachteten Datensatz konnte die Rückfederung beim Biegen von Stahlblechen mit hoher Genauigkeit vorhergesagt werden. Für verschiedene Einsatzszenarien empfehlen sich verschiedene Lernmethoden und für den Gesamterfolg des Einsatzes von maschinellem Lernen erweist sich die Qualität der Trainingsdaten als ausschlaggebend.

**STICHWÖRTER**

Blechbearbeitung, Umformtechnik, Künstliche Intelligenz

**Prediction quality of springback in sheet metal working – A comparative study of machine learning models**

**ABSTRACT** The springback parameter is crucial for the manufacturing accuracy in bending processes. This study examined the suitability of machine learning methods for springback prediction and evaluated the accuracy of the predictions. It was possible to predict springback in bending of steel sheets with high accuracy for a given dataset. It turns out that different use scenarios require different learning methods and that the quality of training data is crucial for overall success when using machine learning methods.

## 1 Motivation

Um konkurrenzfähig zu bleiben, müssen sich Fertigungsunternehmen mit der digitalen Transformation befassen. In kaum einer anderen Branche ist dies für die globale Wettbewerbs- und Zukunftsfähigkeit so wichtig wie im produzierenden Gewerbe. Durch den wachsenden Kostendruck bei gleichzeitig sinkenden Profiten nutzen Unternehmen dieses Sektors zunehmend neue Wertschöpfungsketten und digitale Lösungsstrategien [1]. Auch in der Blechumformung könnte die digitale Transformation einen Beitrag zur Effizienzsteigerung und damit zur Kostenreduktion in der Produktion leisten.

In der umformenden Bearbeitung werden Abweichungen zwischen Soll- und Ist-Geometrie des bearbeiteten Blechs meist durch die Rückfederung nach dem Umformprozess hervorgerufen und müssen zur Wahrung der Maßhaltigkeit kompensiert werden [2]. Kann die Rückfederung schon im Produktentstehungsprozess genau vorhergesagt und damit die benötigte Kompensation bestimmt werden, ist es möglich, die notwendige Nachbearbeitung an der Form oder dem Prozess zu verringern und zeitraubende Anfahrprozesse zu vermeiden. Dadurch kann Zeit- und Materialeinsatz bis zum ersten Gutteil und letztlich Ausschuss reduziert werden [3]. Aktuell werden zur Vorhersage dieser Prozessgröße meistens FEM (Finite-Elemente-Methode)-Simulationen in Verbindung mit empirischen Daten, aber auch rechnerische Abschätzungen genutzt, die helfen, Biegeprozesse zu optimieren [2].

Neben den genannten Verfahren können zur präzisen Vorhersage der Rückfederung auch Methoden des maschinellen Lernens (englisch: Machine Learning, ML) genutzt werden. In diesem Zusammenhang untersuchten Cruz, Barbosa *et al.* die Anwendung eines Neuronalen Netzes (englisch: Artificial Neural Network, ANN) zur Vorhersage der Rückfederung. Trotz des Vorteils einer effizienten Modellierung durch ANNs sollte berücksichtigt werden, dass sie für sich genommen keine explizite Lösung für die Modellierungsaufgabe bieten. Weiterhin wird angegeben, dass andere maschinelle Lernmethoden Vorteile bei der Vorhersage bieten könnten, welche die Schwächen der ANNs ausgleichen. In dieser Studie wurde ein Datensatz verwendet, der mittels realer experimenteller Messungen generiert wurde, was zu einer begrenzten Datenmenge führte [4].

Generell erfordern neuronale Netzwerke für eine zuverlässige Mustererkennung umfangreiche Datensätze. Aus diesem Grund prüft die vorliegende Studie weitere ML-Algorithmen auf deren Korrektheit, Relevanz, Robustheit, Stabilität, ihren Ressourcenverbrauch sowie ihre Interpretierbarkeit.

## 2 Stand der Technik

Um den Umformprozesses, auf dessen Grundlage die Untersuchungen der ML-Algorithmen stattfand, besser zu verstehen, wird im Folgenden auf den Prozess des Freibiegens eingegangen und im Anschluss die Rückfederung erläutert. Weiterhin werden kurz das Maschinelle Lernen und die verwendeten Methoden beschrieben.

## 2.1 Freibiegen

Das Freibiegen ist ein Fertigungsverfahren aus der Gruppe Biegeumformen mit geradliniger Werkzeugbewegung (**Bild 1**).

Bei der Umformung des Halbzeugs bilden die stationären Auflager in Kombination mit dem in einer Achse beweglichen Stempel bei Freibiegevorgängen das Werkzeug. Dieses Werkzeug, der Stempelradius und der Stempelhub sowie die Stärke und die Materialeigenschaften des Blechs haben einen Einfluss auf die erzeugte Formgeometrie. Werkzeug und Werkstück berühren sich im Linienkontakt, während sich die Biegelinie in der Umformzone frei bildet. Der erzeugte Innenradius am geformten Blech ist zudem größer als der eingesetzte Stempelradius des Umformwerkzeugs. Der Stempel erzeugt hier ebenfalls durch eine Translationsbewegung die Umformung der freien Blechgeometrie. Der größte Vorteil des Freibiegens ist die geometrische Flexibilität. Durch den Prozessaufbau ist es durch eine Änderung der Prozessparameter möglich, den Biegewinkel zu verändern, ohne die Biegewerkzeuge austauschen zu müssen. Im Vergleich zu anderen Biegeverfahren (beispielsweise Gesenkbiegen) sind der resultierende Biegewinkel und Biegeradius nach dem Rückfedern des Blechs vor allem aufgrund der Prozessungenauigkeit nur schwer vorhersagbar [5].

## 2.2 Konventionelle Ermittlung der Rückfederung in der Blechumformung

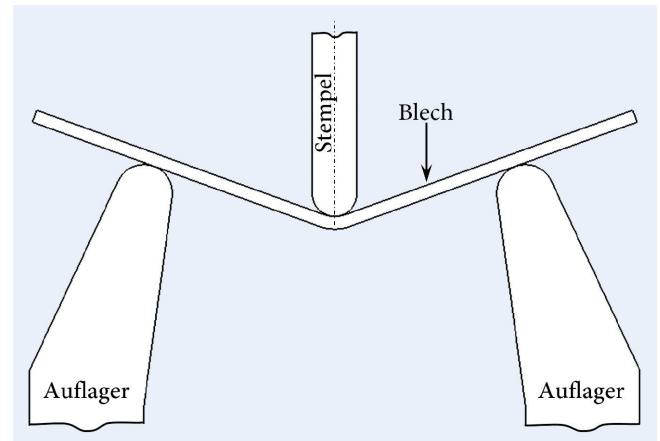
Die Rückfederung (**Bild 2**) ist ein mechanisch unabdingbarer, aber unerwünschter Vorgang beim Biegen von Blechen.

Durch die vom Werkzeug induzierte Umformkraft auf das Blech stellt sich eine elastische oder elastisch-plastische Formänderung ein. Sobald das Blech entlastet wird, stellt sich dieses um den elastischen Anteil der Umformarbeit zurück. Besonders beim freien Biegen ist die Rückfederung uneingeschränkt möglich und wird über die Blecheigenschaften (Härte, Festigkeit), die Stempelgeometrie und die Prozesseigenschaften beeinflusst. Um einen spezifischen Winkel am Blech zu erzeugen, wird die Rückfederung durch Überbiegen kompensiert. Zur rechnerischen Abschätzung der Kompensation sind Kenntnisse der Blechdicke  $s_b$ , des mittleren Biegeradius  $r_m$ , des Elastizitätsmodul  $E$ , der Querkontraktionszahl  $n$ , der Anfangsfließspannung  $k_{f0}$  sowie des elastischen und plastischen Anteils des Biegemoments  $M_{b,elast}$  beziehungsweise  $M_{b,plast}$  notwendig [5]. Zusätzlich zu der bereits aufgeführten Quelle können [2] und [4] herangezogen werden.

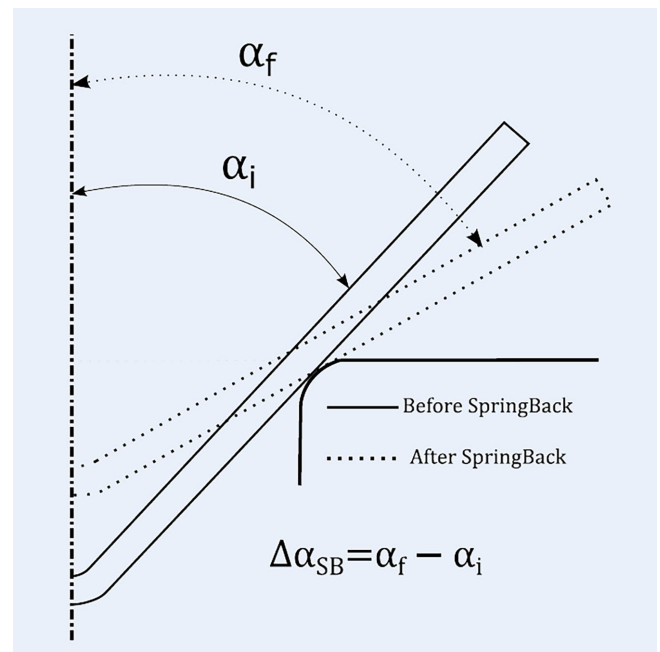
Um die Vorhersage der Rückfederung zukünftig zu vereinfachen, sollen maschinelle Lernprogramme trainiert werden, welche die Kompensation anhand von deutlich weniger Kenngrößen für unterschiedliche Blechgeometrien sowie Materialien vorhersagen können.

## 2.3 Maschinelles Lernen

ML hat das Ziel, durch die Analyse von Daten Wissen zu generieren, um Vorhersagen und Prognosen abzuleiten. Durch das sogenannte Training, ein Prozess, der das Modell mit Eingabedaten und erwarteten Ausgaben versorgt, können ML-Modelle Vorhersagen treffen, ohne dass explizites Programmieren erforderlich ist. [6]. ML-Modelle eignen sich für viele Aufgaben und werden in zwei Hauptkategorien eingeteilt: Regression und Klassifikation. Diese Studie konzentriert sich auf die Vorhersage der



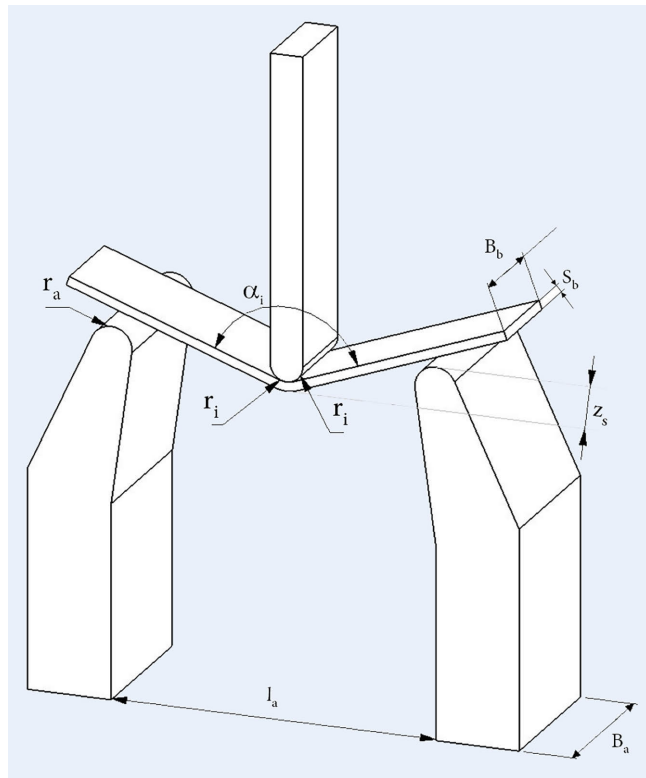
**Bild 1.** Exemplarische Darstellung des Freibiegens. Grafik: Fraunhofer IPA



**Bild 2.** Definition der elastischen Erholung (Rückfederung) beim Biegeprozess. Grafik: [4]

Rückfederung, repräsentiert durch eine kontinuierliche Zahl, und stellt daher ein Regressionsproblem dar [7]. Wichtig für die Modellbewertung und Modellverbesserung ist die richtige Wahl von Metriken, wobei der Mean Absolute Error (MAE) und der Root Mean Squared Error (RMSE) häufig verwendet werden.

In der Studie wurden verschiedene ML-Algorithmen eingesetzt, um die Rückfederung beim Biegen von Blechen vorherzusagen. Die Auswahl der ML-Algorithmen wurde in Anlehnung an Müller und Guido getroffen, welche die populärsten Algorithmen in diesem Bereich nutzen [7]. Aufgrund der Anwendung lag der Fokus auf überwachten Modellen (Supervised Learning). Modelle, die in der Datenerfassung besonders simpel sind, wurden nicht berücksichtigt. Dies schließt beispielsweise lineare Regressionen und einfache Entscheidungsbäume aus. Stattdessen wurden baumbasierte Ensemble-Modelle wie Random Forest, Extra Trees und Gradient Boosting verwendet. Die Auswahl beschränkte sich zudem auf Algorithmen, die sich für das vorliegende Regressionsproblem (die Vorhersage der Rückfederung) besonders eignen,



**Bild 3.** Versuchsaufbau zur Erstellung der Test- und Trainingsdaten für das Training der maschinellen Lernprogramme. Grafik: Fraunhofer IPA

wodurch reine Klassifikationsalgorithmen ausgeschlossen wurden. Ziel der Studie war es, ein möglichst breites Spektrum an Algorithmen zu evaluieren. Daher wurde mit dem Multilayer Perceptron (MLP) lediglich ein neuronales Netzwerkmodell untersucht, ohne komplexere neuronale Modelle einzubeziehen.

Die Methoden umfassen:

- Random Forest (RF): Ein Ensemble-Lernverfahren, das eine Menge an Entscheidungsbäumen während des Trainings generiert und deren Durchschnitt oder Mehrheitsabstimmung für die Vorhersage nutzt.
- Extra Trees (ET): Ähnlich wie Random Forest, aber mit einer höheren Stufe an Randomisierung und Daten, was zu einer effizienteren Berechnung führt.
- Gradient Boosted Trees (GBT): Ein weiteres Ensemble-Lernverfahren, das schwache Entscheidungsbäume in einer sequenziellen Weise trainiert, wobei jeder Baum die Fehler des vorherigen korrigiert.
- Support Vector Machine (SVM): Algorithmus, der durch das Finden der optimalen Trennlinie (oder Hyperfläche in höher-dimensionalen Räumen) arbeitet, kann auch für Regressionsprobleme angepasst werden.
- Multilayer Perceptron (MLP): Ein einfaches neuronales Netzwerk mit einer oder mehreren verborgenen Schichten, das in der Lage ist, nicht-lineare Beziehungen im Datensatz zu modellieren.

### 3 Datenerhebung und Versuchsaufbau

ML-Methoden benötigen für das Training Datensätze, welche für verschiedene Konstellationen der betrachteten Kenngrößen die beobachtete Zielgröße, in diesem Fall die Rückfederung, ent-

halten. Allgemein gesprochen extrahieren sie aus diesen Daten statistische Muster, welche für die Vorhersage der Zielgröße in neuen Datensätzen verwendet werden können. Die Methoden werden anhand von Testdaten evaluiert, für welche die Rückfederung zwar grundsätzlich bekannt ist, aber dem Modell nicht zur Verfügung steht.

Die hier beschriebene Untersuchung bedurfte einer einfachen, aber präzisen Versuchsdurchführung zur Erhebung von validen Trainings- und Testdaten. Dazu wurde der Freibiegeprozess (in Anlehnung an DIN EN ISO 178) im 3-Punkt-Biegeverfahren ausgewählt. Dabei kam für die Biegeversuche eine Universalprüfmaschine der Firma Zwick/Roell vom Typ „Zwick 1554 MOPS“ zum Einsatz. Die Verwendung einer solchen Prüfmaschine bietet sowohl den Vorteil einer reproduzierbaren Prozessführung als auch die Möglichkeit zur präzisen Datenerfassung der benötigten Stempelkraft und des Stempelhubes.

Die variablen Prozessparameter sind der Stempelhub  $z_s$ , der Auflagerabstand  $l_a$  und die Blechstärke  $S_b$ , die bei jedem Versuch zu erfassen sind. **Bild 3** zeigt den Versuchsaufbau mit dem Werkzeug (Auflager und Stempel in Endposition) sowie das gebogene Blech.

Durch Absenken des Stempels in negativer Z-Richtung wird im Versuch der Biegewinkel  $\alpha$  über den Stempelhub indirekt eingestellt. Die definierte Translationsbewegung des Stempels startet am Berührungspunkt von Blech und Stempel. Ist die Endposition erreicht, wird eine Haltezeit von einer Sekunde (maschinenbedingt) abgewartet und anschließend die Richtung der Stempelbewegung umgekehrt.

Die vollständige Entlastung des Blechs ist erreicht, sobald die gemessene Kraft am Stempel auf das Grundniveau absinkt. Der Stempelhub (positive Z-Richtung) zwischen der Endposition und dem vollständigen Abheben des Stempels vom Blech markiert den Bereich der Rückfederung des Blechs. Dieser Definition folgend wird im weiteren Verlauf der eingestellte Biegewinkel und die Rückfederung als Länge in Millimeter und nicht als Winkel in Grad angegeben.

Zur Erstellung des Datensatzes wurden Stahlbleche (1.4301 X5CrNi18–10) und verschiedene Parametereinstellungen gewählt und in den Versuchen in allen möglichen Kombinationen (**Tabelle 1**) dreimal wiederholt durchgeführt. Es ergab sich ein Datensatz mit 403 einzelnen Beobachtungen.

### 4 Methodik

In der Studie wurde ein Vergleich verschiedener ML-Methoden durchgeführt: Die Algorithmen sind Random Forest (RF), Extra Trees (ET), Gradient Boosted Trees (GBT), Support Vector Machine (SVM) und ein Multilayer Perceptron (MLP), die in der Python-Bibliothek von scikit-Learn vorliegen [8].

Die Vergleichskriterien wurden aus der Arbeit von Siebert *et al.* [9] abgeleitet und umfassen Korrektheit, Relevanz, Robustheit, Stabilität und Ressourcenverbrauch. Im Folgenden werden die Kriterien näher erläutert und die jeweils zu ihrer Evaluation genutzten Verfahren genannt. Die Kriterien und Evaluationsmetriken sind zusammengefasst in **Tabelle 2**.

Im Studienkontext sollen ML-Modelle valide sowie präzise Vorhersagen der Rückfederung treffen können. Dabei bezieht sich „korrekt“ auf eine möglichst exakte Vorhersage der Rückfederung durch die ML-Modelle. Die Qualität der Prognose wird anhand der Abweichung zwischen Vorhersage und tatsächlicher

**Tabelle 1.** Konstante sowie variable Versuchsparameter zur Erstellung des Datensatzes.

Versuchsparameter	Formelzeichen	Gewählte Größe					
Blechstärke	$S_b$	0,5 mm	1 mm	1,5 mm	2 mm	2,5 mm	3 mm
Stempelhub	$z_s$		2,5 mm	5 mm	10 mm	15 mm	20 mm
Auflagerabstand	$L_a$		10 mm	20 mm	30 mm	40 mm	50 mm
Stempelradius	$r_s$	0,75 mm					
Originale Blechlänge	$L_b$	80 mm					
Auflagerradius	$r_a$	1 mm					
Auflagerbreite	$B_a$	50 mm					

**Tabelle 2.** Kriterien und Metriken für die Evaluation.

Kriterium	Metrik
Korrektheit	Root Mean Squared Error (RMSE)
Relevanz	Varianz der Kreuzvalidierung (KV), $R^2$
Robustheit	Reaktion auf Rauschen im Datensatz
Stabilität	Leave-One-Out-Cross-Validation (LOOCV)
Ressourceneffizienz	Trainingszeit, Laufzeit

Beobachtung gemessen, wofür der Root Mean Squared Error (RMSE) als Metrik verwendet wird.

Ein relevantes Modell ist in der Lage auch neue, bisher nicht gesehene Daten korrekt vorherzusagen [10]. Dabei ist es nötig, einen optimalen Kompromiss zwischen Bias und Varianz zu finden [9]. Ein Modell mit hohem Bias nimmt starke Annahmen über die Daten vor und neigt dazu, die Muster in den Trainingsdaten zu vereinfachen (underfitting). Ein Modell mit hoher Varianz neigt im Gegenteil dazu, zu viele Details der Trainingsdaten abzubilden (overfitting). In beiden Fällen können die Modelle für ungesehene Daten keine hinreichend genauen Vorhersagen treffen. Der Bias wird mit dem Bestimmtheitswert  $R^2$  zwischen Vorhersage und Beobachtung bewertet. Ein hoher  $R^2$ -Wert bedeutet, dass das Modell einen großen Teil der Varianz erklären kann und wahrscheinlich nicht unter zu hoher Verzerrung leidet. Zur Beurteilung der Varianz wurden die Schwankungen in der Modellqualität für verschiedene Aufteilungen von Trainings- und Testdatensätzen untersucht. Dabei diente die Varianz der Kreuzvalidierung als Metrik. Eine niedrige Varianz in diesem Kontext gilt als Qualitätsindikator.

Ein robustes Modell behält seine Funktionsfähigkeit bei, auch wenn es mit wenigen oder fehlerhaften Daten konfrontiert wird [11]. Typische Fehler sind Ausreißer, fehlende Daten oder Rauschen im Datensatz [12]. Der Leistungsabfall des Modells wurde anhand unterschiedlich stark verrauschter Daten evaluiert, wobei der Root Mean Square Error (RMSE) erneut als Bewertungsmetrik herangezogen wurde.

Ein Modell sollte zudem eine hohe Stabilität aufweisen und konsistente Ergebnisse liefern, wenn es auf verschiedene Teilmengen der Trainingsdaten angewendet wird [9]. Dieser Aspekt gewinnt vor allem dann an Bedeutung, wenn der Trainingsdatensatz, wie in dieser Studie, klein ist. Die Stabilität wurde mit Leave-One-Out-Kreuzvalidierung geprüft und es wurde pro

Datenpunkt ein Modell auf allen übrigen Daten trainiert. Um die Stabilität des Modells zu ermitteln, wurde der durchschnittliche Vorhersagefehler über alle Iterationen berechnet [13]. Ein Modell gilt als stabil, wenn es bei diesem Verfahren konsistente Ergebnisse liefert. Daher wurde die Standardabweichung der Kreuzvalidierungsergebnisse als Indikator für die Stabilität verwendet.

Die Entwicklung und Anwendung eines ML-Modells sollen möglichst ressourcenschonend sein, um das Ziel der Nettoeinsparung von Ressourcen, Materialien sowie Arbeitszeit im Fertigungsprozess zu erreichen. Betrachtet wurde die Laufzeit bei der Vorhersage sowie die Zeit, welche für das Training des Modells benötigt wird [9]. Die Metriken zur Bewertung der Ressourcennutzung umfassen die Trainingszeit sowie die Inferenzzeit in ms. Die Trainingszeit bezieht sich auf die für das Trainieren des Modells benötigte Zeit. Eine kürzere Trainingszeit wird hier bevorzugt, da ein längerer Trainingsprozess mehr Rechnerressourcen erfordert. Die Inferenzzeit bezieht sich auf die Zeit, welche die Modelle für die Vorhersagen benötigen.

## 5 Ergebnisse

**Tabelle 3** präsentiert die Evaluationsergebnisse der ML-Modelle. Die besten Ergebnisse sind hervorgehoben.

Für das Kriterium Korrektheit zeigen alle Modelle einen RMSE von unter 0,25. Das bedeutet, dass die durchschnittliche Modellvorhersage um maximal 0,25 mm vom beobachteten Wert abweicht, entsprechend umgerechnet in Grad beträgt die Abweichung 0,09° bis 0,27°. Die Umrechnung der Rückfederung von Millimetern in Grad ist mittels folgender Formel möglich [14]:

$$\arctan\left(\frac{l_a}{2 \cdot z_s}\right) \times \frac{180}{\pi}$$

Die Beurteilung der Leistungsfähigkeit eines Modells in der Praxis erfordert die Festlegung eines akzeptablen Fehlerspielraums, was in diesem Fall durch die ISO-2786-Norm geschieht. Sie legt verschiedene Genauigkeitsstufen und zugehörige Toleranzen fest.

Die Norm unterscheidet zwischen den Genauigkeitsstufen „f“ (fein), „m“ (mittel), „c“ (grob) und „v“ (sehr grob). Für die in dieser Studie verwendeten Metallteile ist eine Toleranzgrenze von  $\pm 0,5^\circ$  vorgegeben. Diese Grenze bestimmt die maximal zulässige Winkelabweichung für die gebogenen Metallkomponenten. Damit ergibt sich, dass der Vorhersagefehler aller Modelle in die Genauigkeitsstufe „f“ fällt und so höchsten Ansprüchen genügt.

Ein relevantes Modell lernt abstrakte Eigenschaften der Trainingsdaten, ohne sie zu detailliert zu beschreiben. Es sind daher

Tabelle 3. Evaluationsergebnisse

ML-Modell	RMSE	Varianz der KV	R <sup>2</sup>	Reaktion auf Rauschen	Standard-abweichung KV	Trainingszeit (ms)	Inferenzzeit (ms)
Kriterium	Korrektheit	Relevanz	Relevanz	Robustheit	Stabilität	Ressourceneffizienz	Ressourceneffizienz
RF	0,244	0,041	0,784	+	0,047 ± 0,206	24912	2,027
ET	0,231	0,031	0,889	++	0,049 ± 0,194	19541	1,939
GBT	0,242	0,021	0,761	+	0,038 ± 0,183	43688	1,302
SVM	0,200	0,016	0,893	--	0,053 ± 0,244	421007	2,587
MLP	0,201	0,023	0,825	--	0,053 ± 0,244	16466261	17,988

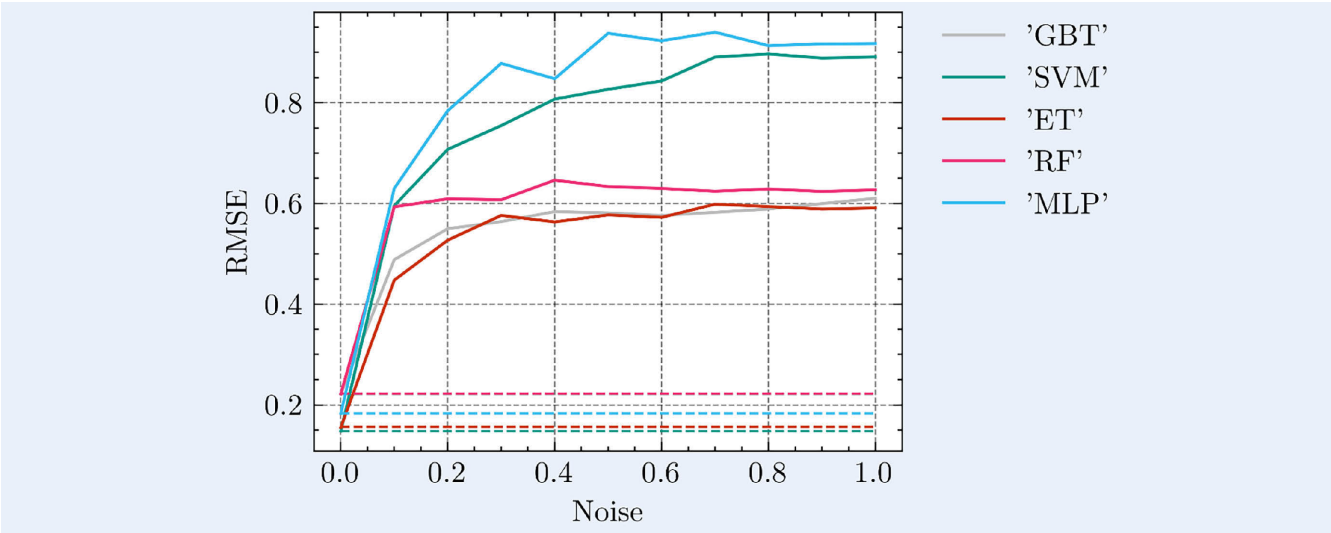


Bild 4. Leistung der Machine-Learning-Modelle mit zunehmend verrauschten Daten. Noise: % hinzugefügtes Rauschen; Root Mean Squared Error RMSE: Vorhersagefehler. Die gestrichelte Linie zeigt die Performanz der Modelle auf den Standarddaten. Grafik: Fraunhofer IPA

mit Varianz der KV und  $R^2$  zwei Maße zur Bewertung nötig. Die Analyse der Ergebnisse aus Tabelle 3 zeigt, dass insbesondere die Modelle SVM und MLP eine niedrige Varianz und hohen  $R^2$  erreichen, was auf hohe Relevanz hinweist.

Die Robustheit der ML-Modelle wurde ursprünglich anhand von zwei Hauptkriterien ermittelt. Erstens, wie die Modelle mit fehlenden Daten umgehen und zweitens, wie die Modelle auf Rauschen im Trainingsdatensatz reagieren. Um das erste Kriterium zu überprüfen, wurden iterativ Daten aus dem Trainingsdatensatz entfernt, um anschließend zu analysieren, welche Auswirkungen dies auf die Leistungsfähigkeit der Modelle hat. Dabei zeigte sich, dass, wenn Daten fehlten, alle Modelle in ähnlichem Maße an Performance einbüßten, sodass keine signifikanten Unterschiede zwischen den Modellen festgestellt werden konnten. Aus diesem Grund ist dieses Kriterium in Tabelle 3 nicht aufgeführt. Für das zweite Kriterium wurde Rauschen zu den Trainingsdaten hinzugefügt, um die Generalisierungsfähigkeiten der Modelle zu beurteilen. Bild 4 zeigt die Leistung der Modelle anhand des RMSE bei zunehmend verrauschten Trainingsdaten.

Die Leistungsfähigkeit aller Modelle nahm rasch ab, als Rauschen in die Daten eingeführt wurde. Bereits bei einer Rausch-

intensität von 20 % war ein signifikanter Anstieg des Vorhersagefehlers zu beobachten. Von allen Modellen zeigte das ET-Modell die robusteste Leistung und übertraf die anderen Modelle. Auffallend ist, dass die bislang besten Modelle MLP und SVM am schlechtesten auf verrauschte Daten reagierten. Eine mögliche Erklärung könnte die Neigung dieser komplexen Modelle zu Overfitting sein, wodurch das Rauschen überproportional in ihre Vorhersagen einfließt. Diese Beobachtungen betonen die Bedeutung einer umsichtigen Modellauswahl sowie einer präzisen Datenerfassung.

Die Stabilität der Modelle wurde durch Betrachtung der Standardabweichung des Fehlers bei der Leave-One-Out-Kreuzvalidierung beurteilt. Ein Modell mit einer geringen Standardabweichung über alle Teilmengen hinweg deutet auf ein stabiles Modell hin. Das GBT-Modell ist in dieser Hinsicht das stabilste. Ähnlich wie bei der Robustheit weisen die MLP- und SVM-Modelle im Vergleich zu den anderen Modellen geringere Stabilitätsniveaus auf, obwohl sie bei der Bewertung der Korrektheit am vielversprechendsten waren.

Die Ressourcennutzung wird in dieser Studie in Bezug auf die Laufzeit der Modelle bewertet. Aus den Ergebnissen wird ersicht-



lich, dass die Modelle ET und RF die kürzeste Trainingszeit haben. Das MLP-Modell ist mit über 16 Sekunden deutlich das langsamste. Ein vergleichbares Bild ergibt sich für die Inferenzzeit. Dieses Mal ist das GBT-Modell das schnellste Modell, während das MLP weiterhin das Langsamste ist.

## 6 Fazit

Die Evaluation verschiedener ML-Modelle für die Vorhersage der Rückfederung beim Freibiegen zeigt, dass sehr genaue Vorhersagen möglich sind, die den Normanforderungen im höchsten Maße genügen. Allerdings schneidet kein ML-Modell in allen Kriterien gleichermaßen gut ab. Daher sollten je nach Anwendungsfall in der Praxis andere Modelle ausgewählt werden. Im Folgenden sind einige definierte Anwendungsfälle aufgeführt:

In Industrien, in denen mit hoher Präzision gearbeitet werden muss, ist das Kriterium Korrektheit am wichtigsten. Ein Beispiel für solch hochpräzise Metallkomponenten wäre die Luft- und Raumfahrtindustrie. Hier empfehlen sich Modelle wie SVM und MLP, welche die Genauigkeit priorisieren.

Wird mit einer Vielzahl verschiedener Metallkomponenten gearbeitet, müssen ML-Modelle an diese Bedingungen anpassbar sein. Hier empfiehlt sich ein Modell, das sowohl relevant als auch robust ist, um eine gleichbleibende Leistung über verschiedene Projekte hinweg zu gewährleisten. In Bezug auf die Relevanz gehörten das SVM- und MLP-Modell zu den besten Modellen. Allerdings zeigten gerade diese beiden Modelle Defizite in der Robustheit, vor allem wenn sie mit verrauschten Daten konfrontiert wurden. Das ET-Modell ist in der Kategorie Relevanz durchschnittlich, besticht jedoch durch hohe Robustheit. Daher wird es insbesondere für Anwendungen mit variablen Daten empfohlen.

In einigen Industrien ist eine variierende Datenqualität zu erwarten, weshalb robuste Modelle für wechselnde Datenqualität benötigt werden. Hier eignen sich Modelle, die gut mit Rauschen im Datensatz umgehen können. Die Ergebnisse dieser Studie zeigen, dass im Kriterium Robustheit das ET-Modell die beste Leistung erzielt.

Unternehmen mit begrenzten technischen Ressourcen, die Biegeteile im kleinen Maßstab herstellen, können mittels ML den Biegeprozess optimieren, ohne viele Ressourcen zu verbrauchen. Die Ergebnisse zeigen, dass sich hier alle getesteten ML-Modelle mit Ausnahme des MLP-Modells eignen. Das MLP ist ein komplexes Modell, das sowohl lange Trainings- als auch Inferenzzeit erfordert.

Ein zweites praxisrelevantes Ergebnis der Studie ist die klare Abhängigkeit der Performanz der ML-Modelle von der Qualität der Trainingsdaten. Durch Beimischung von Rauschen zu den mit hoher Genauigkeit erhobenen Studiendaten steigt der Vorhersagefehler aller betrachteten Modelle stark an. Es ist für die Nutzung eines ML-Modells zur Optimierung der Produktionsprozesse also ausschlaggebend, dass Trainingsdaten mit ausreichender Genauigkeit erhoben werden.

## Literatur

- [1] Stahl, B.: Digital Transformation in the Manufacturing Sector – Tackling Business and Technology Challenges. Dissertation, Universität Bayreuth, 2023
- [2] Hoffmann, H.; Neugebauer, R.; Spur, G. (Hrsg.): Handbuch Umformen. München: Carl Hanser Verlag 2012
- [3] Gössling, M. J.: Metamodell unterstützte Simulation und Kompensation von Rückfederungen in der Blechumformung. Aachen: Shaker Verlag 2010
- [4] Cruz, D. J.; Barbosa, M. R.; Santos, A. D. et al.: Application of Machine Learning to Bending Processes and Material Identification. Metals 11 (2021) 9, #1418
- [5] Mertin, C. F.: Experimentelle und numerische Analyse von mehrstufigen Stanzbiegeprozessen für Bauteile mit federnden Eigenschaften. Dissertation, Rheinisch-Westfälische Technische Hochschule Aachen, 2019
- [6] El Naqa, I.; Murphy, M. J.: What Is Machine Learning? In: El Naqa, I.; Li, R.; Murphy, M. (eds.): Machine Learning in Radiation Oncology. Cham: Springer 2015, pp. 3–11, doi.org/10.1007/978-3-319-18305-3\_1
- [7] Müller, A.; Guido, S.: Introduction to Machine Learning with Python: A Guide for Data Scientists. Sebastopol, CA: O'Reilly Media 2016
- [8] Pedregosa, F.; Varoquaux, G.; Gramfortet, A. al.: Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research 12 (2011), pp. 2825–2830
- [9] Siebert, J.; Joeckel, L.; Heidrich, J. et al.: Construction of a quality model for machine learning systems. Software Quality Journal 30 (2022) 2, pp. 307–335
- [10] Zhou, Z.-H.: Machine Learning. Singapore: Springer Singapore 2021
- [11] Liu, X.; Du, Y.; Lu, X. et al. (eds.): 2019 6th International Conference on Frontiers of Industrial Engineering (ICFIE 2019). Proceedings. London: Institute of Electrical and Electronics Engineers (IEEE) 2019
- [12] Sáez, J. A.; Luengo, J.; Herrera, F.: Evaluating the classifier behavior with noisy data considering performance and robustness: The Equalized Loss of Accuracy measure. Neurocomputing 176 (2016), pp. 26–35
- [13] James, G.; Witten, D.; Hastie, T. et al.: An Introduction to Statistical Learning. New York: Springer US 2021
- [14] Kurrle, P.: Predicting the Spring Back in Sheet Metal Forming. A Comparative Study of Machine Learning Models using Design Science Research. Master Thesis, Hochschule für Technik Stuttgart, 2023, noch nicht publiziert



**Peter Lange, M.Sc.**

Foto: Autor

**Philipp Kurrle, M.Sc.**

Fraunhofer-Institut für Produktionstechnik und Automatisierung IPA  
Nobelstr. 12, 70569 Stuttgart  
Tel. +49 711 / 970-1468  
peter.lange@ipa.fraunhofer.de  
www.ipa.fraunhofer.de

**Prof. Dr. Ulrike Pado**

Hochschule für Technik Stuttgart  
Schellingstr. 24, 70174 Stuttgart  
Tel. +49 711 / 89262811  
ulrike.pado@hft-stuttgart.de  
www.hft-stuttgart.de

## LIZENZ



Dieser Fachaufsatz steht unter der Lizenz Creative Commons Namensnennung 4.0 International (CC BY 4.0)

