

# Machine Learning Applications in Sheet Metal Constitutive Modelling: a review

**Armando E. Marques<sup>a, e, \*</sup>, Tomás G. Parreira<sup>a, e</sup>, André F. G. Pereira<sup>a, e</sup>, Bernardete M. Ribeiro<sup>b</sup> and Pedro A. Prates<sup>a, c, d, e</sup>**

<sup>a</sup> CEMMPRE, Department of Mechanical Engineering, University of Coimbra, 3030-788 Coimbra, Portugal

<sup>b</sup> CISUC, Department of Informatics Engineering, University of Coimbra, 3030-290 Coimbra, Portugal

<sup>c</sup> TEMA, Centre for Mechanical Technology and Automation, Department of Mechanical Engineering, University of Aveiro, 3810-193 Aveiro, Portugal

<sup>d</sup> LASI, Intelligent Systems Associate Laboratory, 4800-058 Guimarães, Portugal

<sup>e</sup> ARISE, Advanced Production and Intelligent Systems Associated Laboratory, 4200-465 Porto, Portugal

\* Correspondence: armando.marques@uc.pt

## **Abstract**

The numerical simulation of sheet metal forming processes depends on the accuracy of the constitutive model used to represent the mechanical behaviour of the materials. The formulation of these constitutive models, as well as their calibration process, has been an ongoing subject of research. In recent years, there has been a special focus on the application of data-driven techniques, namely Machine Learning, to address some of the difficulties of constitutive modelling. This review explores different methodologies for the application of Machine Learning algorithms to sheet metal constitutive modelling. These methodologies include the use of machine learning algorithms in the identification of constitutive model parameters and the replacement of the constitutive model by a metamodel created by a machine learning algorithm. A discussion about the merits and limitations of the different methodologies is presented, as well as the identification of some possible gaps in the literature that represent opportunities for future research.

**Keywords:** Machine Learning, Constitutive Modelling, Parameter Identification, Metamodeling, Data-driven Learning, Sheet Metal Forming

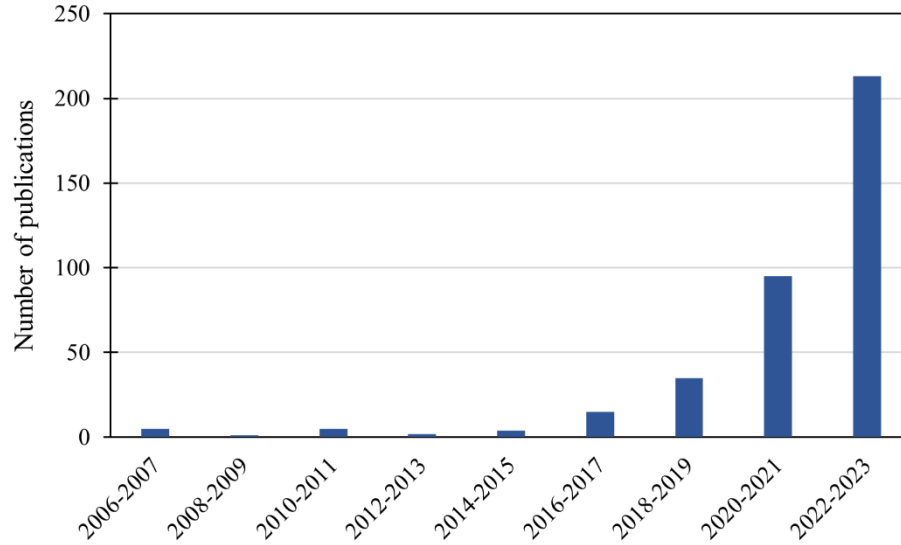
## 1. Introduction

Sheet metal forming is a manufacturing process that is widely used in the automotive, metalworking and aerospace industries, due to its capability to produce large quantities of components at a relatively low cost (Marques et al., 2020). These industries operate in a highly competitive market, where continuous innovation is key to maintaining profits while ensuring that quality, safety, and environmental requirements are met. In this context, a trial-and-error approach to process design and optimisation proves to be unfeasible, as it leads to significant costs, in terms of time, process redesign and scrap losses (Diller et al., 1997). Finite Element Method (FEM) simulation proves to be a more interesting alternative, as it can be used to design and optimise sheet metal forming processes, even though they are characterized by high non-linearity in the boundary conditions, geometry and material properties. FEM results rely on the accurate representation of the mechanical behaviour of metal sheets, which is usually described by constitutive models (Yoshida et al., 2013). Detailed overviews of sheet metal constitutive modelling include the works of Hou et al. (Hou et al., 2023) and Banabic et al. (Banabic et al., 2020).

Constitutive models typically contain several material parameters, which are traditionally identified from the results of a large number of standardized mechanical tests (Habraken et al., 2022). Classical mechanical tests, such as the uniaxial tensile and simple shear tests, allow simple strain paths with quasi-homogeneous deformation in the region of interest. However, the plastic behaviour of a material during a forming process generally involves heterogeneous deformation with strain path changes under multiaxial loads. Also, the development of new metallic alloys has led to the development of more flexible constitutive models, which contain more parameters to be identified, and consequently, require more homogeneous mechanical tests (Rabahallah et al., 2009). To overcome these limitations, alternative strategies to identify constitutive model parameters have been proposed. These include inverse strategies, that make use of heterogeneous mechanical tests that allow to characterize the plastic behaviour of the material for a wide range of stress and strain states, usually by taking advantage of full field measurement techniques, such as Digital Image Correlation (DIC) (Prates et al., 2016; Rossi et al., 2022). Some of the most widely used inverse identification techniques (Avril et al., 2008) include Finite Element Model Updating (FEMU) (Prates et al., 2016), Constitutive Equation Gap Method (CEGM) (Latourte et al., 2008), Virtual Fields Method (VFM) (Fu et al., 2017) and Equilibrium Gap Method (Périé et al., 2009). In these techniques the optimization problem is generally solved through derivative-free

or gradient-based algorithms, which can converge to non-optimal solutions (local minima). Genetic-based algorithms can be used instead, which can find the optimal solution, however, they are associated with high computational costs (Chaparro et al., 2008).

In recent years, with the increase in available data and computational power, data-driven approaches, such as Machine Learning (ML), have been in the forefront of research for many different areas of science. Jin et al. (Jin et al., 2023) presented a review focussed on the application of ML in various areas of solid mechanics, including fracture mechanics, biomechanics and architected materials, while the work of Brodnik et al. (Brodnik et al., 2023), focussed in the application of ML to various areas of experimental solid mechanics and the works of Li et al. (Li et al., 2024) and Dornheim et al. (Dornheim et al., 2024) present general overviews of ML applications to the problem of constitutive modelling. In the context of sheet metal forming, these approaches have been successfully applied to a variety of problems, including springback prediction (Jamli et al., 2015; Ur et al., 2021), blank force optimisation (Feng et al., 2019), defect detection (Jingdong et al., 2017; Nath and Chattopadhyay, 2021), fracture initiation (X. Li et al., 2023; Müller et al., 2021), manufacturing process optimisation (Cruz et al., 2021; Liu et al., 2021) and process control (Senn et al., 2014). Given the versatility of these approaches, their application to sheet metal constitutive modelling has been presented as a promising alternative to the current solutions. For instance, when compared to the inverse identification techniques, they are less sensitive to the problem of local minima. In this context, some significant progress has already been made, particularly in the last few years, even though the number of published works can still be considered low. For example, a simple search in Google Scholar using the terms “machine learning”, “sheet metal” and “constitutive model” leads to around 393 published works. Figure 1 shows the biannual distribution of these results, where a clear ascending trend in the number of published works can be observed, particularly since 2016.



**Fig.1** Yearly distribution of references in the results from Google Scholar when searching for “machine learning”, “sheet metal” and “constitutive model”.

Although Machine Learning (ML) has been at the forefront of several areas of sheet metal forming research, its application to sheet metal constitutive modelling is still limited. In this work, different applications of ML algorithms to the problem of sheet metal constitutive modelling, focused primarily on material plastic behaviour, are reviewed and discussed.

The aim is to promote the dissemination of the state-of-the-art ML applications in sheet metal forming problems, in particular the use of ML in identification strategies and constitutive modelling of the material plastic behaviour, as well as to identify the future research trends. This work will cover the two main methodologies based on the application of ML algorithms, resorting to (i) using ML to identify constitutive model parameters and (ii) using ML to create metamodels that replace the constitutive models. To the authors knowledge, no other review work in this area has included the previously mentioned methodologies in the context of sheet metal constitutive forming.

The following chapter consists of a brief overview of machine learning, including a description of the different categories of ML algorithms. In the third chapter, the various methodologies and applications of ML algorithms from the literature are identified and discussed. The fourth chapter identifies possible areas of future study, and the final chapter addresses the final conclusions of this work.

## 2. Machine Learning

Machine Learning (ML) is an area of study focused on developing systems capable of learning without the need for being explicitly programmed (Jang et al., 2021). ML

algorithms are particularly useful for solving complex tasks that would require the creation of a first-principle model, as the implementation of such a model can be an unfeasibly expensive process regarding both resources and time. These algorithms can be grouped into three main categories, based on the configuration of the data they require: (i) supervised learning algorithms, (ii) unsupervised learning algorithms, and (iii) reinforcement learning algorithms.

Supervised learning algorithms are ML algorithms that learn based on available labelled data, i.e., the dataset contains both the inputs and outputs (labels) of the problem. This learning process is commonly referred to as training and consists of mapping the relationships between input and output variables present in the data provided to the algorithm. Based on these relationships, the internal parameters of a model are adjusted, allowing it to predict the outputs for new sets of unseen input data (Zhou, 2018). Supervised algorithms have been used to solve various types of problems, including classification (e.g. (Nath and Chattopadhyay, 2021)), regression (e.g. (Ur et al., 2021)), image recognition (e.g. (Aljuaid and Anwar, 2022; Buhuş et al., 2017)) and text identification (e.g. (Budhiraja and Mago, 2020)).

Unsupervised learning algorithms are trained with unlabelled data, and their purpose is to identify patterns present within the dataset (Ghahramani, 2004). These algorithms can be used, for example, in clustering problems where the number of classes in the dataset is not defined, which can lead in a posterior phase to the identification of previously unforeseen, or undesirable classes. Some interesting applications of unsupervised learning in engineering problems include the works of Stanev et al. (Stanev et al., 2018), where ML is used to process X-ray diffraction data, Brito et al. (Brito et al., 2021), where unsupervised learning is used to detect faults on bearings, and Kunka et al. (Kunka et al., 2021), Generale et al. (Generale et al., 2024) and Muth et al. (Muth et al., 2023), which all use Principal Component Analysis as a dimensionality reduction tool.

Reinforcement learning algorithms are focused on the idea of training an intelligent agent that makes decisions in a problem environment, to maximize a measure of cumulative reward (Martín-Guerrero and Lamata, 2021). The problems to be solved by reinforcement learning are typically situations where choices with various options are presented, in succession, and the goal of the agent is to find the choice in each case that leads to an overall better result. The works of Villarreal et al. (Villarreal et al., 2023) and Wang et al. (Wang et al., 2021) are examples of the successful application of reinforcement learning. The first work present a methodology to generate designs of experiments for material model calibrations, while the second presents an innovative, non-cooperative game, for the calibration of material constitutive laws.

The various applications explored in this work use supervised learning algorithms, since the problems solved consist of regression problems. As such, the following sections contain a brief theoretical presentation of the most commonly used supervised learning algorithms. The advantages and drawback of each ML algorithm, when applied to regression problems, are summarized in Table 1.

Algorithm	Advantages	Drawbacks
Multi-Layer Perceptron	<ul style="list-style-type: none"> <li>• Good general predictive performance;</li> <li>• High flexibility.</li> </ul>	<ul style="list-style-type: none"> <li>• Low interpretability;</li> <li>• Performance depends on choice of architecture (number of hidden layers and nodes per layer).</li> </ul>
Recurrent Neural Networks and Long Short-Term Memory networks	<ul style="list-style-type: none"> <li>• Good general predictive performance;</li> <li>• Capable of exploiting sequential data.</li> </ul>	<ul style="list-style-type: none"> <li>• Low interpretability;</li> <li>• Computationally expensive;</li> <li>• Vanishing and exploding gradient problems during training (Recurrent Neural Network).</li> </ul>
Convolutional Neural Networks	<ul style="list-style-type: none"> <li>• Good general predictive performance;</li> <li>• Suitable for large datasets;</li> <li>• Capable of reducing problem dimensionality without loss of information.</li> </ul>	<ul style="list-style-type: none"> <li>• Low interpretability;</li> <li>• Computationally expensive;</li> <li>• Requires high amounts of data.</li> </ul>
Gaussian Processes	<ul style="list-style-type: none"> <li>• Good general predictive performance;</li> <li>• Capable of generalizing with small amounts of data;</li> <li>• Not very sensitive to noisy data;</li> <li>• Provides measurements for prediction uncertainty.</li> </ul>	<ul style="list-style-type: none"> <li>• Low interpretability;</li> <li>• Not suitable for large datasets;</li> <li>• High memory usage.</li> </ul>
Decision Trees and other tree-based algorithms	<ul style="list-style-type: none"> <li>• Good interpretability, particularly for single Decision Trees.</li> </ul>	<ul style="list-style-type: none"> <li>• Poor performances due to the non-continuity of the solutions;</li> <li>• High memory usage for algorithms with many trees.</li> </ul>
Support Vector Machines	<ul style="list-style-type: none"> <li>• Good general predictive performance;</li> <li>• Capable of generalizing with small amounts of data;</li> <li>• Reduced influence of outliers.</li> </ul>	<ul style="list-style-type: none"> <li>• Low interpretability;</li> <li>• High memory usage;</li> <li>• Very sensitive to kernel choice.</li> </ul>
Symbolic Regression	<ul style="list-style-type: none"> <li>• Good interpretability;</li> <li>• User can control the complexity of the equations generated.</li> </ul>	<ul style="list-style-type: none"> <li>• If the equation generated is not properly restricted, interpretability can be diminished.</li> </ul>

Table 1. Advantages and drawbacks of the different ML algorithms.

## 2.1 Multi-Layer Perceptron (MLP)

The Multi-Layer Perceptron (MLP) is one of the simplest, and most commonly used types of neural networks, applicable to both classification and regression problems. Often referred to simply as an Artificial Neural Network (ANN), it is a feed forward network, composed of multiple layers of nodes (neurons) (Murtagh, 1991). Each node in a given layer connects to the nodes in the next layer, passing information, but never connects with other nodes in the same layer. The input layer is the first layer of an MLP metamodel and contains as many nodes as there are input features in the dataset. It is followed by a user specified number of layers called hidden layers. The number of hidden layers, as well as the number of nodes in each hidden layer are defined before the model training process. Based on metamodel performance, these values can be optimized, which requires the training of a new metamodel. In the last layer, referred to as output layer, predictions are made based on the information received from the previous layers. A schematic representation of an MLP, with two hidden layers, can be found in Figure 2 a).

The MLP algorithm is applicable to multi-output problems natively, with the number of nodes in the output layer equalling the number of output variables of the problem. The output of a single node from a hidden layer is given by:

$$z_i = \emptyset(\sum_j w_{ij} z'_j + d_i), \quad (1)$$

where  $z_i$  is the output of node  $i$  from the current layer,  $z'_j$  is the output of node  $j$  from the previous layer,  $w_{ij}$  is the weight associated with  $z'_j$ ,  $d_i$  is a bias term and  $\emptyset$  is a non-linear activation function. For classification problems, the output layer nodes have a similar formulation to the hidden layer nodes, while for regression problems, the output layer nodes may not have an activation function, which is the same as using an identity activation function.

The training of the MLP neural network metamodel consists of adjusting the weights in order to learn the mapping function from inputs to outputs, with the so-called backpropagation algorithm (Dreyfus, 1990). This is done by assessing whether each weight value should be increased or decreased, and then changing all the weights in the network accordingly by a small increment, in an iterative process until a minimum prediction error is achieved (Dreyfus, 1990; Murtagh, 1991).

This algorithm possesses high flexibility, and as such is suitable for complex problems, but suffers from low interpretability.

## **2.2 Recurrent Neural Networks (RNN)**

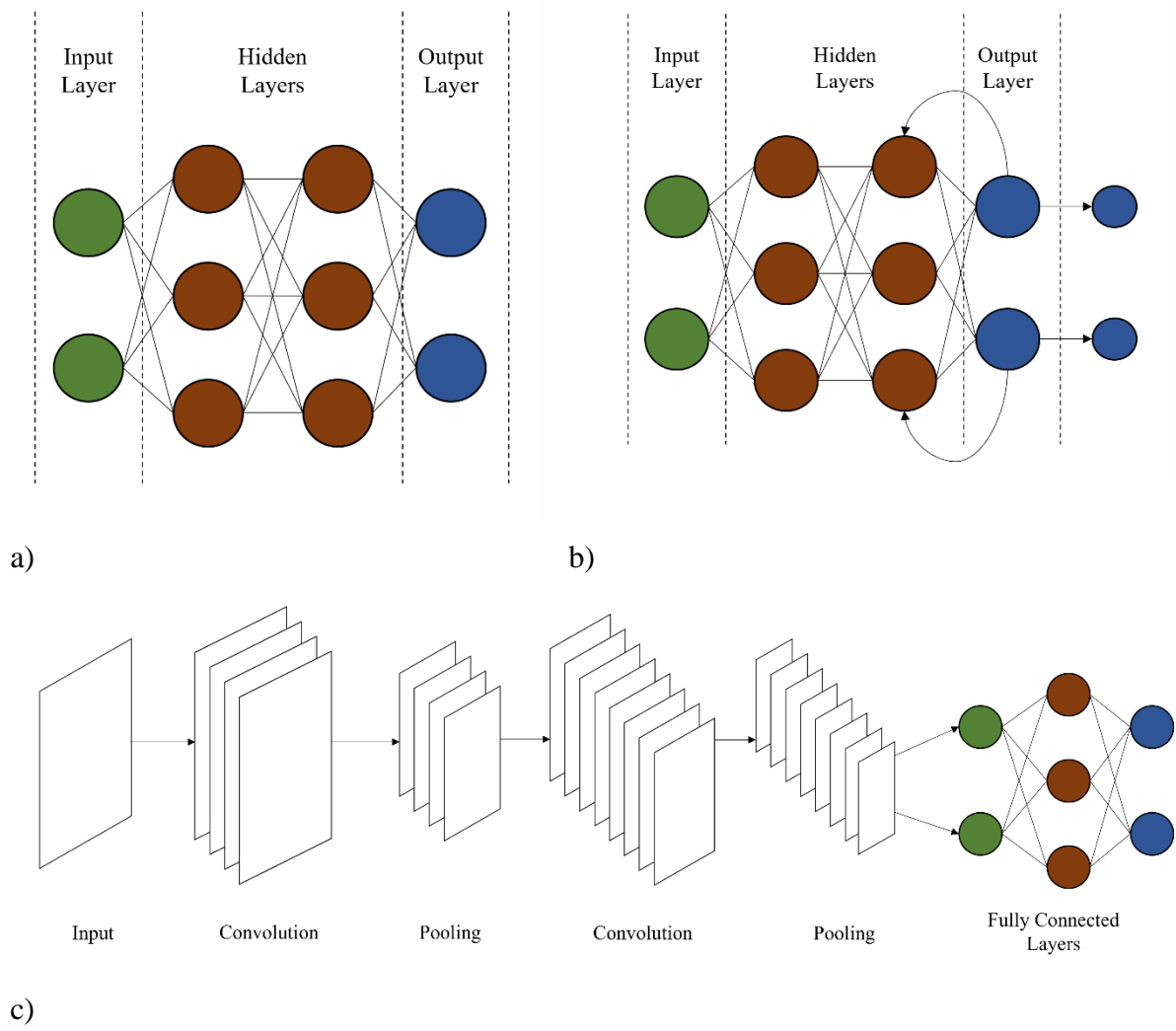
Recurrent Neural Networks (RNN) are neural networks with the capability to interpret sequential data, and as such are well suited to applications where the inputs are not independent of each other, but instead have sequential relationships that can be exploited (Sherstinsky, 2020). The capability to interpret sequential data is the main reason for RNNs to be used in problems with associated time series (Nketiah et al., 2023; Rosenkranz et al., 2023). A schematic representation of a simple RNN is shown in Figure 2 b). Traditional RNN have a problem with vanishing or exploding gradients during training, since they tend to have a large depth. In order to take this into account some variations have been developed, with the most commonly used one being the Long Short-Term Memory (LSTM) neural network. These algorithms suffer from low interpretability, and are computationally expensive (Salehinejad et al., 2017).

## **2.3 Convolutional Neural Networks (CNN)**

Convolutional Neural Networks (CNN) are a type of neural network designed primarily to take images as inputs, although they can also take 1D data. The data is processed by the algorithm to identify features and patterns, assigning importance (weights) to them. These neural networks have three different types of layers: convolutional layers, pooling layers and fully-connected layers (Albawi et al., 2018). The convolutional layers are responsible for assigning weights, while the pooling layers apply a dimensionality reduction to the problem, reducing the number of inputs. A CNN can have any combination of convolutional and pooling layers, as long as the first layer is always a convolutional layer. Similarly to the hidden layers in an MLP, the number of convolutional and pooling layers is defined before the training process. This number depends on the problem to be solved by the CNN, and while there is no rule that governs this choice, more complex problems generally require more layers. The fully connected layers are the last layers. They receive the inputs from the previous layers, in vector form, and make the final predictions. The fully connected layers are similar to the layers of an MLP metamodel. A schematic representation of a CNN can be found in Figure 2 c).

This algorithm is suitable for large datasets, since it can reduce problem dimensionality, without significant loss of information. However, it tends to require large amounts of data to achieve good problem generalization, is computationally expensive, and like other neural network algorithms, suffers from low interpretability (Gu et al., 2018).





**Fig.2** Schematic representations of different types of neural network configurations: a) MLP; b) RNN; c) CNN.

## 2.4 Gaussian Process (GP)

A Gaussian Process (GP) is a collection of random variables that follows a Gaussian distribution and are fully defined by a mean function and a covariance function (Jingdong et al., 2017). Generally, the mean function is assumed to be zero, so the covariance function is enough to completely define the GP. The GP regression metamodel can be represented by:

$$y(\mathbf{x}) = f(\mathbf{x}) + \epsilon, \quad (2)$$

where  $y(\mathbf{x})$  is an observed output,  $f(\mathbf{x})$  is the corresponding GP variable and  $\epsilon$  represents a zero mean Gaussian white noise. Assuming that  $\mathbf{y}(\mathbf{x}^t)$  is the target vector of outputs present in the training dataset, and  $\mathbf{y}(\mathbf{x}^p)$  is the vector of outputs to be predicted, the joint normal probability distribution is given by:

$$\begin{bmatrix} \mathbf{y}(\mathbf{x}^t) \\ \mathbf{y}(\mathbf{x}^p) \end{bmatrix} \sim N \left( 0, \begin{bmatrix} \mathbf{K}(\mathbf{X}, \mathbf{X}) + \sigma_\epsilon^2 \mathbf{I} & \mathbf{K}(\mathbf{X}, \mathbf{X}_*) \\ \mathbf{K}(\mathbf{X}_*, \mathbf{X}) & \mathbf{K}(\mathbf{X}_*, \mathbf{X}_*) \end{bmatrix} \right), \quad (3)$$

with  $\sigma_\epsilon^2$  being the noise variance,  $\mathbf{I}$  being the identity matrix and each  $\mathbf{K}$  being a covariance matrix evaluated for all considered points, with  $\mathbf{X}$  representing the data from the training dataset, and  $\mathbf{X}_*$  the unseen data for which the metamodel will make predictions. Finally, the GP metamodel predictions are given by the following equations:

$$\bar{\mathbf{f}}_* = \mathbf{K}(\mathbf{X}_*, \mathbf{X})[\mathbf{K}(\mathbf{X}, \mathbf{X}) + \sigma_\epsilon^2 \mathbf{I}]^{-1} \mathbf{y}(\mathbf{x}^t), \quad (4)$$

$$\mathbf{cov}(\mathbf{f}_*) = \mathbf{K}(\mathbf{X}_*, \mathbf{X}_*) - \mathbf{K}(\mathbf{X}_*, \mathbf{X})[\mathbf{K}(\mathbf{X}, \mathbf{X}) + \sigma_\epsilon^2 \mathbf{I}]^{-1} \mathbf{K}(\mathbf{X}, \mathbf{X}_*), \quad (5)$$

where  $\bar{\mathbf{f}}_*$  is the vector of predictions (mean), and  $\mathbf{cov}(\mathbf{f}_*)$  represents the covariance of the metamodel outputs, which acts as a measure of prediction uncertainty.

This algorithm can generalize a problem with relatively small amounts of data and is robust against noise. However, it has low interpretability, and due to memory requirements, is not suitable for application to large datasets (Liu et al., 2014). The GP algorithm is used primarily for regression problems (Wang et al., 2018; Zhang and Xu, 2022), although there are examples of classification applications (Saunders et al., 2022).

## 2.5 Decision trees (DT) and Tree-based algorithms

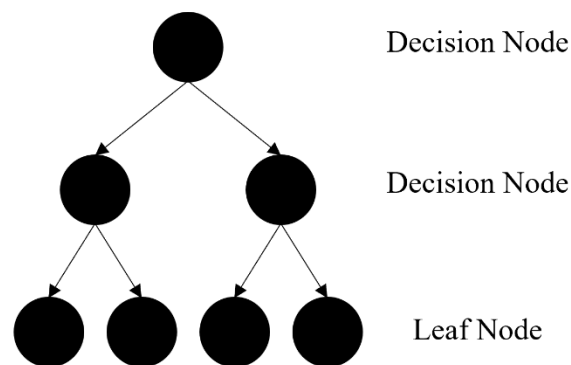
The Decision Tree (DT) algorithm works by applying a continuous splitting process to the training data, based on simple rules, that are chosen to minimize an error metric in the resulting nodes (Rokach and Maimon, 2005). A node where this split occurs is called a decision node. A schematic representation of a DT is shown in Figure 3. The most common error metric used on this case is the MSE (mean square error). The splitting process is repeated until each of the resulting final nodes has an MSE below a pre-defined threshold, or until all final nodes have less than pre-defined number of training data points. These final nodes, responsible for the predictions of the model, are usually called leaf nodes. The predictions from each node correspond to the average of the output values from the training points in the node.

A Random Forest (RF) algorithm generates an ensemble metamodel consisting of a series of decision trees, each trained with a randomly generated sample from the training data (Breiman L., 2001). The predictions made by each decision tree are compiled, and the average value is chosen as the final prediction of the RF metamodel.

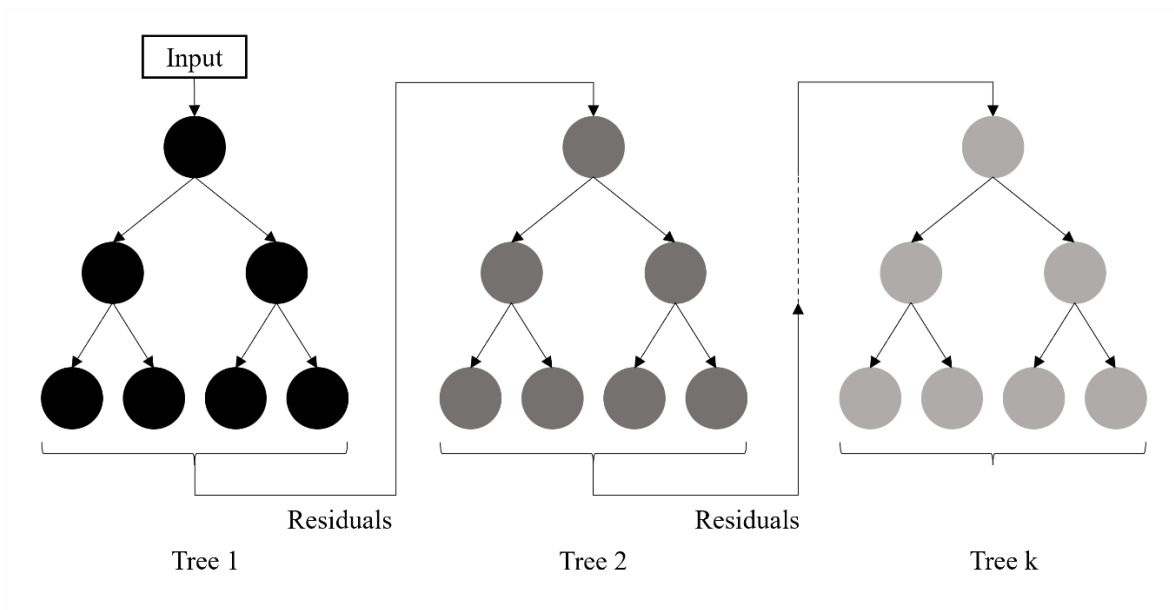
The Gradient Boosting Decision Trees (GBDT) is an extension of the RF algorithm. In this algorithm, the trees are created sequentially, with each subsequent tree including the residuals of the previous tree in the training data. (Friedman, 2001). A schematic

representation of a GBDT metamodel can be found in Figure 4. In this way, each tree attempts to minimize the error of the previous tree. Like with the RF algorithm, the final prediction of the ensemble metamodel is the average of the predictions made by each tree in the metamodel.

One significant advantage of the DT algorithm, and to a lesser extent also an advantage of the other tree-based algorithms, is the fact that it is possible to observe the rules used to split the training data. This means that the tree-based metamodels are interpretable, unlike the metamodels created by most other algorithms, which are usually considered black-box models. However, due to the non-continuous nature of the solutions, these algorithms do not present good predictive performances for complex regression problems. Also, if the trained metamodel has a considerable number of trees, the memory requirements of the metamodel can become prohibitive.



**Fig.3** Schematic representation of a Decision tree.



**Fig.4** Schematic representation of a Gradient Boosting Decision Trees ensemble metamodel with  $k$  trees.

## 2.6 Support Vector Machines (SVM)

The Support Vector Machines (SVM) algorithm, when applied to regression problems, fits a function to the available data, while keeping the function as linear as possible to avoid overfitting. This is achieved by considering an error value  $\gamma$ , under which errors are accepted without penalty (Smola and Schölkopf, 2004). In practice, this means finding the function that can encompass the most training data points in the area around it, with a distance of  $\gamma$  or less. To give some flexibility to the metamodel, soft margins can be defined, in the form of slack variables  $\xi_i$  and  $\xi_i^*$ . Points at a distance between these variables and  $\gamma$  still affect the shape of the function, but with a penalty. When considering a linear problem, this metamodel is given by:

$$\begin{cases} \min \left( \frac{1}{2} \|\mathbf{w}\|^2 + V \sum_i (\xi_i + \xi_i^*) \right) \\ \text{with:} \\ y_i - wx_i - \beta_0 \leq \gamma + \xi_i \\ wx_i + \beta_0 - y_i \leq \gamma + \xi_i^* \end{cases}, \quad (6)$$

where  $\mathbf{w}$  is the normal weight vector to the surface of the approximated function and  $V$  is a constant representing the trade-off between tolerance for deviations above  $\gamma$  and function flatness.

For non-linear problems, this algorithm can be generalized by introducing the kernel trick (Schölkopf, 2000). A kernel consists of a similarity function between the inputs of the training data and the inputs for which the metamodel will make predictions. The kernel trick transforms the data to a higher dimensional space, allowing a linear metamodel to learn non-linear functions without a specific mapping. The SVM algorithm is typically capable of generalizing a problem with small amounts of data and is not significantly influenced by outliers in the training data. However, it has low interpretability, suffers from high memory usage, and the predictive performance of the metamodels created by this algorithm is heavily influenced by the choice of kernel (Sangeetha and Kalpana, 2010).

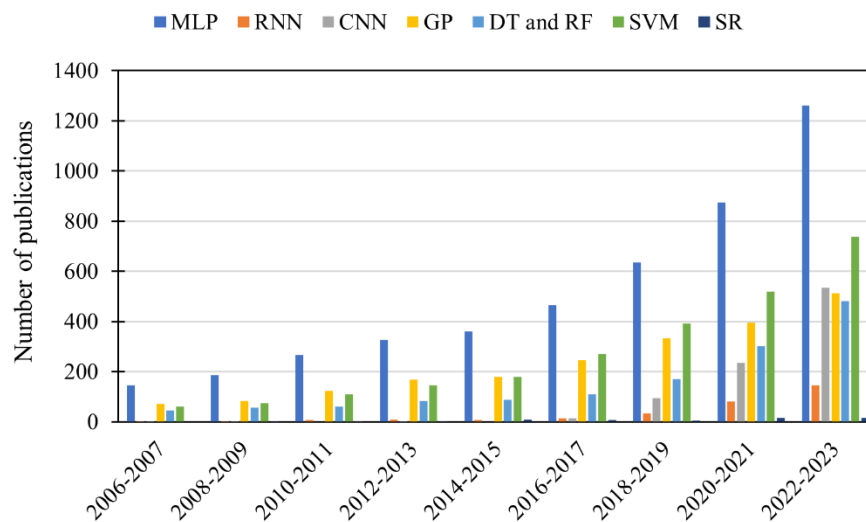
## 2.7 Symbolic Regression (SR)

The Symbolic Regression (SR) algorithm generates a mathematical function to fit the available data. Usually, this algorithm is implemented with Genetic Programming concepts, resulting in an iterative process in which a number of possible equations is generated and evaluated (Augusto and Barbosa, 2000). Then, the equations that better fit the data are recombined into a new set of equations, which are evaluated again until a suitable final equation is found. Like the tree-based algorithms, the metamodels (equations) generated by this

algorithm are usually easy to interpret, but this depends on the format of the generated equation. The resulting equation can be restricted by defining what mathematical operators are allowed, as well as a maximum length for the equation (Angelis et al., 2023). If the algorithm is not properly restricted, the resulting equation may be too long, and depend on too many variables, to be interpretable.

## 2.8 Applications of ML algorithms to sheet metal forming problems

The biannual distribution of ML algorithm applications across the various sheet metal forming problems is presented in Figure 5. This figure presents the number of publications found in Google Scholar when searching for “sheet metal” and the name of each ML algorithm.



**Fig.5** Yearly distribution of references in the results from Google Scholar when searching for “sheet metal” and the various machine learning algorithms.

This distribution shows the clear preference toward the application of MLP, particularly when compared to the other neural network algorithms, which only recently started to be used (especially CNNs). The GP, SVM and tree-based algorithms have also been used for many applications. In contrast, the SR algorithm is virtually unused.

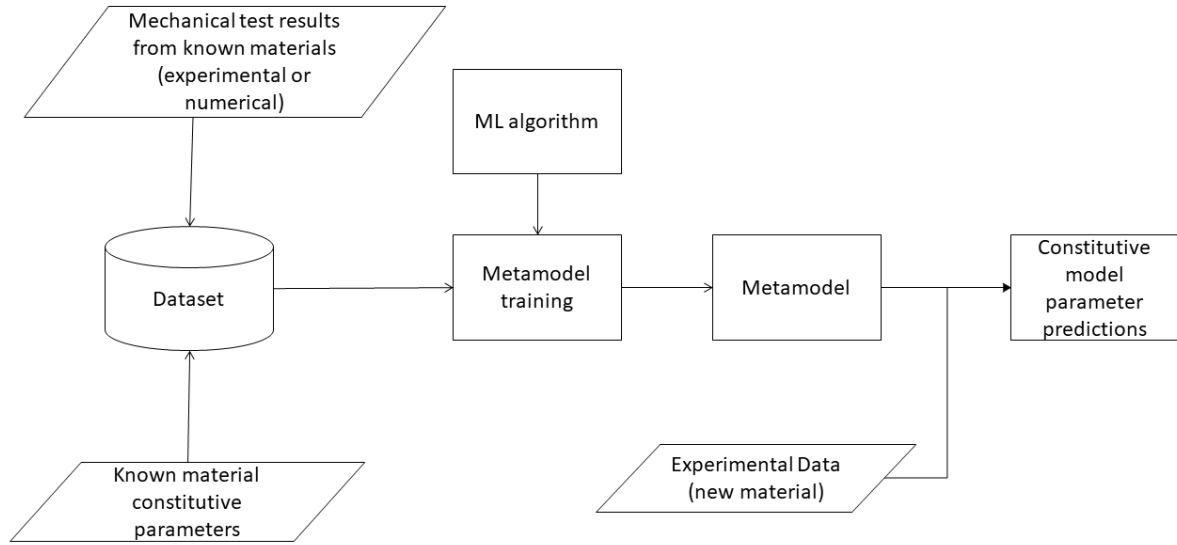
## 3. Review of ML applications in constitutive modelling

Several works have been published involving the use of ML algorithms in the field of constitutive modelling. Such ML applications usually follow one of two main methodologies: (i) ML algorithms are used to identify the parameters for a specific constitutive model formulation (Aguir et al., 2011; Cruz et al., 2021; Guo et al., 2021; Huang et al., 2017; Morand and Helm, 2019; Yao et al., 2021), and (ii) ML algorithms are used to create

metamodels that completely replace the constitutive model (Fazily and Yoon, 2023; Gorji et al., 2020; Hartmaier, 2020; Heidenreich and Mohr, 2024; Jang et al., 2021; Kabliman et al., 2021; Q. J. Li et al., 2023; Pandya et al., 2020; Park et al., 2021; Pham et al., 2022; Zhang and Mohr, 2020). In the first methodology, the ML algorithms create a metamodel that surrogates a specific process, usually a mechanical test, from which the parameter identification is performed, while in the second methodology, the metamodels created surrogate the constitutive model for arbitrary processes, usually with the purpose of implementing the metamodel into FEM code. These methodologies are explored in the following sections.

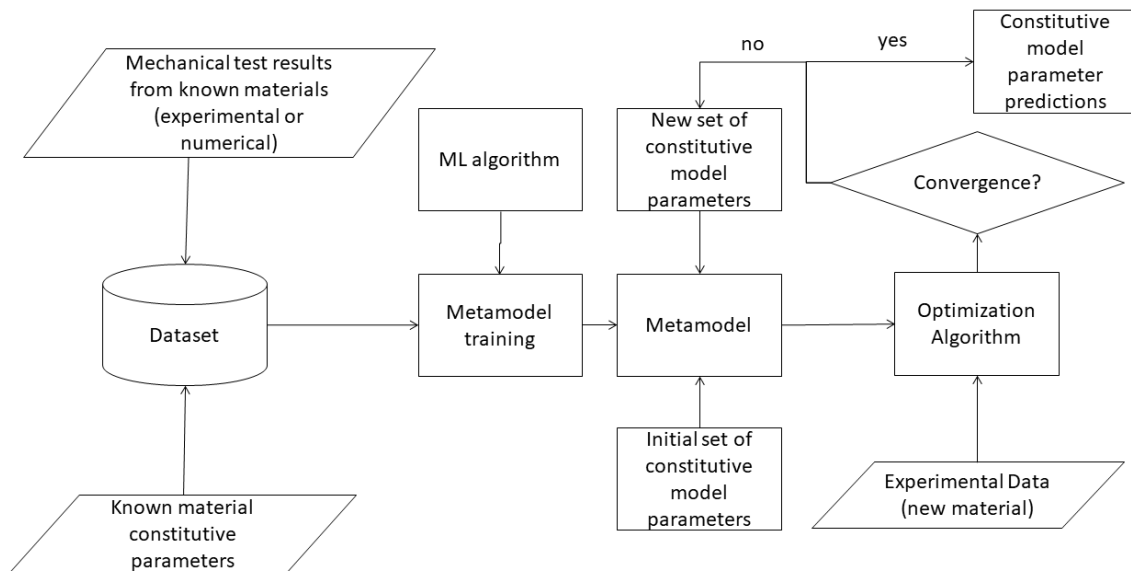
### **3.1 Applications to constitutive parameter identification**

The application of ML algorithms to the identification of constitutive model parameters is an inverse identification method in which the optimization process includes the training of a metamodel. In this training process, optimization algorithms are used to obtain the optimal set of internal metamodel parameters. The constitutive model parameter identification can typically be achieved in one of two ways. In the first approach, the ML metamodel is used to solve the inverse identification problem directly, in which case, the chosen inputs correspond to the data that is available for new materials, at the time of the identification (usually results from mechanical tests, such as forces, displacements, and strains), while the outputs are the constitutive model parameters. The experimental results, obtained from the material to be characterised, are then introduced into the trained metamodel, to predict the constitutive parameters. A simple schematic representation of this approach is shown in Figure 6. Examples of this approach, explored later in this section, include the works of Cruz et al. (Cruz et al., 2021), and Morand et al. (Morand and Helm, 2019).



**Fig.6** Example of the constitutive parameter identification workflow, with a ML metamodel solving the identification problem directly.

The second approach consists of pairing the ML metamodel with classical optimization algorithms. In this case, the metamodel usually takes the constitutive model parameters as inputs and the mechanical test results as outputs. In this context, the optimization algorithm is then responsible for finding a set of constitutive parameters for which the mechanical test results match the outputs of the ML metamodel. This approach is represented schematically in Figure 7. An example of this approach, explored later in this section, is the work of Aguir et al. (Aguir et al., 2011).



**Fig.7** Example of the constitutive parameter identification workflow, considering a pairing of a ML metamodel with an optimization algorithm.

In the following paragraphs, some of the applications of ML to constitutive parameter identification found in the literature are examined.

Cruz et al. (Cruz et al., 2021) identified the Swift hardening law (Swift, 1952) parameters, using an ANN trained with numerical data from the punch force-displacement curves obtained from the simulation of three-point bending tests. The simulations considered parameters of the hardening law typical of various steel grades, and the approach was tested for a dual-phase steel DP590. Huang et al. (Huang et al., 2017) used kriging to identify the Swift hardening law parameters as well as the parameters of a Lemaitre damage model (Bouchard et al., 2011), based on data obtained from the numerical simulation of uniaxial tensile tests of an AA2017 T4 aluminium. In this work, the metamodel prediction is then used to perform a numerical simulation, which is compared with the experimental results. Here, “kriging” is a term frequently used in some areas of study to refer to a particular case of a Gaussian Process, typically used in low dimensional spaces. These works represent applications of a single ML algorithm to the identification of constitutive model parameters.

While the results obtained by these approaches are generally positive, some researchers have sought more complex methodologies. The works of Aguir et al. (Aguir et al., 2011) and Yao et al. (Yao et al., 2021) are such examples. In both works, the use of Neural Networks was combined with a genetic algorithm (D’Angelo and Palmieri, 2021). In the first work, the authors identified the parameters of the Karafillis and Boyce yield criterion (Karafillis and Boyce, 1993) and the Voce isotropic hardening law (Voce, 1948), for an AISI 304 stainless steel. This approach consists of using two ANNs to predict the displacements of pre-selected points, in a circular bulge test and a plane strain tensile test, based on the parameters to be identified. Then, a genetic algorithm is responsible for the optimisation, determining if the constitutive parameters used as input for the ANNs are acceptable, and if not, defining a new set of parameters to be given as input to the ANNs, in an iterative process. The initial parameters considered were obtained experimentally from a uniaxial tensile test, and the datasets used to train the ANNs contained numerical simulation results, of the two mechanical tests mentioned previously. After the initial results, further validation of the methodology was performed for a third test, an elliptical bulge test. The second work consisted of the identification of the parameters of a viscoplastic-damage constitutive model. ANNs were used to determine a suitable initial population for the application of the genetic algorithm. The experimental data used to train the neural networks consisted of stress-strain curves, obtained from uniaxial tensile tests for different strain rates and temperatures. The methodology was



then tested with experimental results from an AA6061 aluminium and achieved good predictive performance.

Another example is the work of Guo et al. (Guo et al., 2021), in which a deep learning metamodel was developed to identify constitutive parameters, based on the combination of a convolutional neural network, used to filter noise in the input data, with a long-short term memory neural network (LSTM). In this context, deep learning is a term usually applied to a neural network based metamodel with multiple layers between the input and output layers. This metamodel was used to identify the elastic parameters (Young's modulus and Poisson's ratio) and three parameters of an exponential hardening law, based on the results of a uniaxial tensile test, including the strain fields from a rectangular region of the sample, load history and geometry. The training was performed with numerical data. After initial tests, the proposed metamodel proved to be sufficiently robust by maintaining low prediction errors when using noisy data. Further validation was performed by comparing the performance of the proposed metamodel with that of a FEMU method, when applied to experimental uniaxial tensile test results of an AA6061 aluminium. The predictions obtained by both methods were in agreement. The authors emphasised the difference in efficiency between the two methods, with the ML method taking significantly longer to train and obtain a prediction than the FEMU method takes to identify the parameters for one material. However, subsequent predictions by the ML method are almost instantaneous, whereas the FEMU method has to repeat the process for each individual material, making the ML method much more efficient after training.

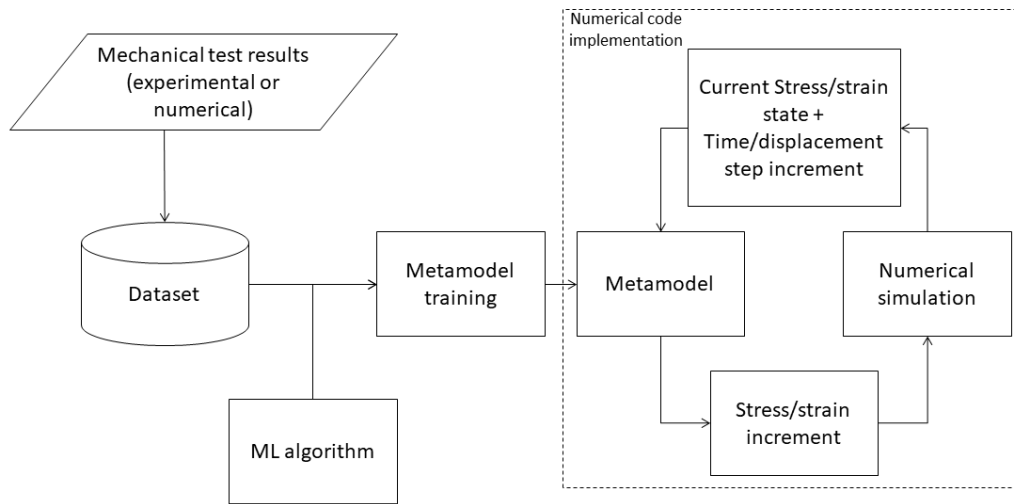
A limitation of the application of a single ML algorithm, identified by Morand et al. (Morand and Helm, 2019), is the ability to properly generalize a problem with non-unique solutions. In this work, the authors considered a hardening law that can have multiple sets of parameters representing the same mechanical behaviour (multiple solutions). For this problem, an ensemble metamodel comprised of multiple neural networks was used, and it achieved a much better performance than the single ANN it was compared to. In this case, both training and testing were performed with numerical data. The ensemble method was then further validated, with the prediction of the hardening law parameters for three experimentally tested materials, an AA6014 aluminium, a DC04 steel and a DP600 steel. This experimental data was extracted from uniaxial tensile tests. The authors concluded that a single ANN metamodel should not be used in this situation, because it tends to average the potential parameter sets, resulting in low predictive performance.

Another limitation of these approaches is the need to pre-select a constitutive model. As the ML metamodels are only used to identify the constitutive parameters, the overall accuracy of the results still depends on the precision of the pre-selected constitutive model to describe the elastoplastic mechanical behaviour of the materials.

A summary of the various works presented in this section can be found in Table 2.

### 3.2 Constitutive model replacement with machine learning metamodels

The approaches presented in the previous section have limitations inherent to the constitutive model formulations. To overcome them, some researchers have explored the potential of using a ML metamodel to completely replace the constitutive model, since the ML algorithms that generate these metamodels are more flexible. In this approach, the metamodel maps the relationship between the stress and strain states, as shown in Figure 8.



**Fig.8** Implementation of ML based metamodels in numerical simulations, as replacement for constitutive models.

Zhang and Mohr (Zhang and Mohr, 2020) and Jang et al. (Jang et al., 2021) both used ANNs to create metamodels that replace von Mises plasticity (von Mises, 1913) and isotropic hardening, based on numerical data, specifically, the first work used results of a uniaxial tension-compression test, and the second work used results of single element simulation for different stress paths (uniaxial tension, simple shear, plane stress and tension-compression cyclic loading). The resulting metamodels performed well and were successfully integrated into FEM code. In the first work, the metamodels were validated in the simulation of a notched uniaxial tensile test, while in the second work, simulations of a dog-bone uniaxial tensile test, and a circular cup drawing were performed, as well as single element simulations in the stress paths used for training. Pandya et al. (Pandya et al., 2020) characterized the plasticity and

fracture response of AA7075 aluminium, considering the influence of temperature and strain rate, based on the results of four different mechanical tests (Uniaxial tension test, notched specimen tension test, central hole tension test and in-plane shear test). This characterization was carried out using a modified Johnson-Cook plasticity model (G.R.Johnson and W.H.Cook, 1983), and a neural network metamodel (MLP). The training data included both numerical and experimental data. For validation, both approaches were implemented in FEM code, and the numerical results obtained for the mechanical tests mentioned previously were compared to experimental results. While the first approach achieved good results only for the pre-necking phase of the tests, the neural network metamodel was satisfactory for both pre-necking and post necking.

Park et al. (Park et al., 2021) used a Radial Basis Function neural network, which is equivalent to an ANN using the radial basis function as the activation function, to model the thermal-plastic behaviour of a DP980 steel. The metamodel was trained with experimental stress-strain data and integrated into a numerical code and applied to the simulation of a warm forming process (V-bending), achieving good accuracy. Gorji et al. (Gorji et al., 2020) used RNNs and MLPs to model the plasticity of an AA5182 and a DC05 steel, when considering strain path changes. All metamodels were trained with numerical data. Multiple problems were studied, including uniaxial tension with loading direction change, radial loading and arbitrary loading with plane stress conditions. Validation was performed through the application of the constitutive model to a 2D foam, and the results highlighted the potential of using ML algorithms to model plasticity. However, the authors also concluded that, in practice, this type of application still has some limitations that must be overcome, particularly in obtaining the required data. The authors argue that for this type of approach to be feasible, hundreds of different stress-strain curves would need to be experimentally measured for a given material.

Fazily and Yoon (Fazily and Yoon, 2023) applied MLPs to the modelling of plasticity behaviour, considering two different training sets based on numerical data from single element simulations considering various stress paths. The first training set is based on an AA6022, modelled with the Hill'48 (Hill, 1948) criterion and the second one is based on a Type 409 stainless steel, modelled with the Yld2000-2d (Barlat et al., 2003) criterion. The generated metamodels take the plane stress tensor of a given increment and the previous flow stress as inputs, and give the updated plane stress tensor as output. These metamodels were integrated into an ABAQUS code and tested, first on a single element subjected to a uniaxial tensile load and then in the simulation of a cup drawing process. Once again, the results of the simulations that used the MLP metamodels were able to match the results obtained with the

yield criteria previously mentioned, while representing a 17% reduction in computation time. Li et al. (Q. J. Li et al., 2023) proposed a more complex methodology, named encoder-decoder framework, which essentially consists of using two different models to make predictions. The first model, called the encoder, receives the problem inputs and projects them into a lower dimensional space. Then, the second model, named decoder, receives the lower dimensional data as input, and outputs the final prediction. Multiple algorithms were used for the encoder, including Convolutional Neural Networks, Recurrent Neural Networks and a transformer (Vaswani et al., 2017), while the decoder consists of a simple fully connected neural network (MLP). In this case, two datasets, containing experimental data consisting of stress-strain pairs were considered. The first one was obtained from uniaxial tensile tests of an unspecified aluminium alloy, considering both monotonic loading and loading-unloading cycles, and the other considered different strain rates and temperatures, for an AISI 316L steel. The developed metamodels were validated for unseen loading scenarios, and overall, achieved excellent predictive performance in interpolation cases, and acceptable performance in extrapolation cases. The metamodel using a transformer encoder had the best performance. Pham et al. (Pham et al., 2022) also applied MLP algorithms to identify the flow curve of an AA5052-H32 aluminium. In this work, both the constitutive model parameter identification approach (by identifying the parameters of a hardening law), and the replacement approach were explored and compared. Four metamodels were generated, two for each approach, with the difference between them being the inclusion of the initial yield stress in the inputs. The training data is numerical, obtained from V-shape tests. The metamodels were validated by simulating a uniaxial tensile test and a truncated cone, and comparing their results with experimental results. Overall, the metamodels from both approaches achieved good representations of the flow curves, with the metamodels that included the initial yield stress performing better, since the other metamodels tended to overestimate the stresses for low plastic strains.

Heidenreich and Mohr (Heidenreich and Mohr, 2024) proposed an alternative approach, based on the use of RNNs. The concept of a common core is introduced, which consists of a set of internal metamodel parameters that are identical for all similar materials. By applying this concept, it is possible to significantly reduce training times when including new materials in the dataset, while at the same time achieving better generalization when compared to metamodels trained in the traditional way, with all internal parameters varying. The metamodels in this work were trained from a numerical dataset, based on von Mises materials with hardening described by a Swift + Voce hardening law, considering random strain paths. Validation of the methodology was carried out through the numerical simulation

of a circular cup drawing, where the proposed methodology surpassed the performance of RNNs trained in the traditional way.

Although the majority of the approaches used various types of neural networks to replace the constitutive model, a few researchers proposed the use of other types of machine learning algorithms. Hartmaier (Hartmaier, 2020) proposed a methodology to generate the yield function based on a SVM algorithm. The algorithm is trained with numerical data representing the stress states, and whether these stress states lead to an elastic or an elastic-plastic deformation, for a reference virtual material with elastic properties typical of a steel and Hill-like anisotropy. The resulting metamodels were applied in a FEM code and validated with the simulations of simple loading conditions (uniaxial stress in two directions, bi-axial strain under plain stress and pure shear), with satisfactory results. Kabliman et al. (Kabliman et al., 2021) used SR metamodels to estimate three parameters of a mean dislocation density model (MD<sup>2</sup>M) constitutive model (Mecking and Kocks, 1981) and to replace the model completely, based on experimental data from hot compression tests from an AA6082 aluminium. These approaches were validated with the simulation of the same test. When compared to the traditionally calibrated physics-based model, the performance of the metamodels was inferior. However, the authors defend that the SR metamodels are more flexible. This is because the physics-based model is only applicable for the temperature/strain rate ranges that were observed during calibration, while the SR metamodels are mathematical expressions that are, theoretically, applicable to any temperature/strain rate combination.

A summary of the various works presented in this section can be found in Table 3.

### **3.3 Discussion**

The analysis of current ML algorithm applications shows the clear trend towards the use of neural networks, as shown in Table 4. This focus on neural networks highlights their capability and versatility but means that the potential of other algorithms is not properly explored. In fact, from the works previously presented, there is only one application of the SVM algorithm, one application of a Symbolic Regression and one application of a Gaussian Process (in the form of kriging). These algorithms have already been successfully used multiple times in many different applications (Asadzadeh et al., 2021; Bomarito et al., 2021; Kubik et al., 2022; Wang et al., 2018; Yu-Lin et al., 2013; Zhang and Xu, 2022), and as such, they represent interesting alternatives to neural network algorithms that warrant further study in the area of sheet metal constitutive modelling.

ML algorithm		Papers	Total
Neural networks	MLP	Cruz et al., 2021, Aguir et al., 2011, Yao et al., 2021, Morand and Helm, 2019, Zhang and Mohr, 2020, Jang et al., 2021, Pandya et al., 2020, Park et al., 2021, Gorji et al., 2020, Fazily and Yoon, 2023, Li et al., 2023, Pham et al., 2022	14
	MLP ensemble	Morand and Helm, 2019	
	RNN/LSTM	Guo et al., 2021, Gorji et al., 2020, Li et al., 2023, Heidenreich and Mohr, 2024	
	CNN	Guo et al., 2021, Li et al., 2023	
SVM		Hartmaier, 2020	1
GP (kriging)		Huang et al., 2017	1
SR		Kablman et al., 2021	1

Table 4. ML algorithms present in the examined papers, for the parameter identification and constitutive model replacement with metamodel approaches.

The applications of ML algorithms to identify constitutive model parameters presented in this work focus on constitutive model formulations with a relatively small number of model parameters to identify, with the exception of the work of Yao et al. (Yao et al., 2021). Also, in the presented cases, all parameters to be identified are continuous, which makes the identification problem a pure regression problem. However, some of the more complex constitutive models include parameters that are not continuous, such as the CPB'06 yield criterion (Cazacu et al., 2006), which includes a single parameter that must be an integer. The identification of these non-continuous parameters may prove a challenge for the regression ML algorithms typically used. Most works presented for this type of application used numerical data for training purposes, but among them there are cases of validation with experimental data, or data with artificial noise, that show the robustness of these approaches (Guo et al., 2021; Morand and Helm, 2019).

Another aspect to consider, is the type of data used to train the metamodels. In most of the works presented in this review, global data, such as force/displacement curves, is used, obtained from tests with low diversity of stress paths (tension-compression (Cruz et al., 2021), uniaxial tension (Huang et al., 2017; Yao et al., 2021), biaxial tension and plane strain (Aguir et al., 2011)). While this data proved sufficient for the applications in study, it would be interesting to consider more applications of full-fields data, with heterogeneous tests (such as the cruciform biaxial tensile test (Prates et al., 2016)). The work of Guo et al. (Guo et al., 2021)

does make use of full-fields data, and the approach presented does successfully deal with some of the challenges this data presents (volume of data, noise), by applying a CNN that is responsible for initial data transformation and denoising. One other challenge associated with full-fields data is the risk of data drop-out during measurements. In work of Hamel et al. (Hamel et al., 2023), the authors analysed the impact of data drop-out on the performance of a physics-informed neural network, and concluded that the methodology was reasonably resistant to this problem.

The application of ML metamodels to replace constitutive models has a clear limitation, highlighted by the works presented, in the fact that the data used to train the metamodels is material dependent, meaning the resulting metamodels are only applicable for the same type of material that the training data was originally obtained from. As such, it can be concluded that while this type of application can provide more modelling flexibility than the traditional phenomenological constitutive models, and as such potentially better accuracy, the metamodels created are specialized for a specific material. This means that a new metamodel must be trained for each different material type considered. Also, the increased flexibility of the algorithms means that the resulting metamodels are more sensitive to the quality of the data used to train them than the metamodels used in the constitutive parameter identification methodology. The presence of noise in the training data can lead to instability when incorporating the metamodel in FEM, especially when different stress/strain paths are poorly represented. As such, for the metamodel to be reliable, the data used to train them must be representative of many different stress/strain paths, and reasonably free of noise. Most of the presented works focus on the use of numerical data, however, in some of the works presented (Kabliman et al., 2021; Pandya et al., 2020; Park et al., 2021) experimental data was successfully incorporated in the training process of the metamodels, showing that this limitation can be surpassed.

Another possible limitation of these approaches comes from the fact that the yield surface convexity is not guaranteed. No steps were taken to guarantee convexity of the surfaces described by the ML metamodel in the presented works, but Hartmaier (Hartmaier, 2020) does identify this limitation, suggesting that this is an area that demands future research work. In fact, this limitation is not restricted to this type of approach. In the applications of ML to the identification of constitutive model parameters, the same problem can occur, if the constitutive model formulation does not guarantee convexity for all combinations of parameters. This is the case, for example, of the yield criterion proposed by Cazacu and Barlat (Cazacu and Barlat, 2001). In this situation, the convexity must be tested for each predicted set of parameters.

The advantages and drawbacks of each approach are summarized in Table 5.

Methodology	Advantages	Drawbacks
Constitutive parameter identification	<ul style="list-style-type: none"> <li>• The metamodel is applicable to different materials (dependent only on the flexibility of the constitutive model).</li> <li>• Small amount of data.</li> </ul>	<ul style="list-style-type: none"> <li>• Performance is affected by the chosen constitutive model.</li> <li>• Information loss during metamodel training.</li> </ul>
Constitutive model replacement with metamodel	<ul style="list-style-type: none"> <li>• Not limited by constitutive model formulations (more flexibility in terms of the materials that can be described).</li> </ul>	<ul style="list-style-type: none"> <li>• Each metamodel is only applicable to the type of material it was trained for.</li> <li>• Moderate risk of instabilities resulting from quantity and quality of the training data.</li> <li>• Information loss during metamodel training.</li> </ul>

Table 5. Advantages and drawbacks of the different ML application approaches.

#### 4. Future work

From the works presented previously, multiple avenues for future research can be identified. First, the application of ML algorithms to the identification of constitutive model parameters can be extended to more complex constitutive models (Cazacu et al., 2006; Plunkett et al., 2008). These models present new challenges, not only because they are associated with a large set of parameters to be identified, but also because the set is composed of discrete and continuous parameters, making the identification process more complex. Secondly, the identification of constitutive model parameters can also be extended to kinematic hardening law parameters (Lemaitre and Chaboche, 1990), for which the deformation history plays an important role and where metamodels with the ability to interpret sequential data may prove to be more accurate (Rosenkranz et al., 2023). Applications of ML to the identification of kinematic hardening law parameters have been successful before (Huber and Tsakmakis, 1999a, 1999b), but not in the context of metal sheets, to the authors knowledge.

Besides these areas of study, other emerging field of research present opportunities and new approaches that can be expanded to the field of sheet metal constitutive modelling (Lourengo et al., 2022). Some of these fields are briefly explored in the following sections.



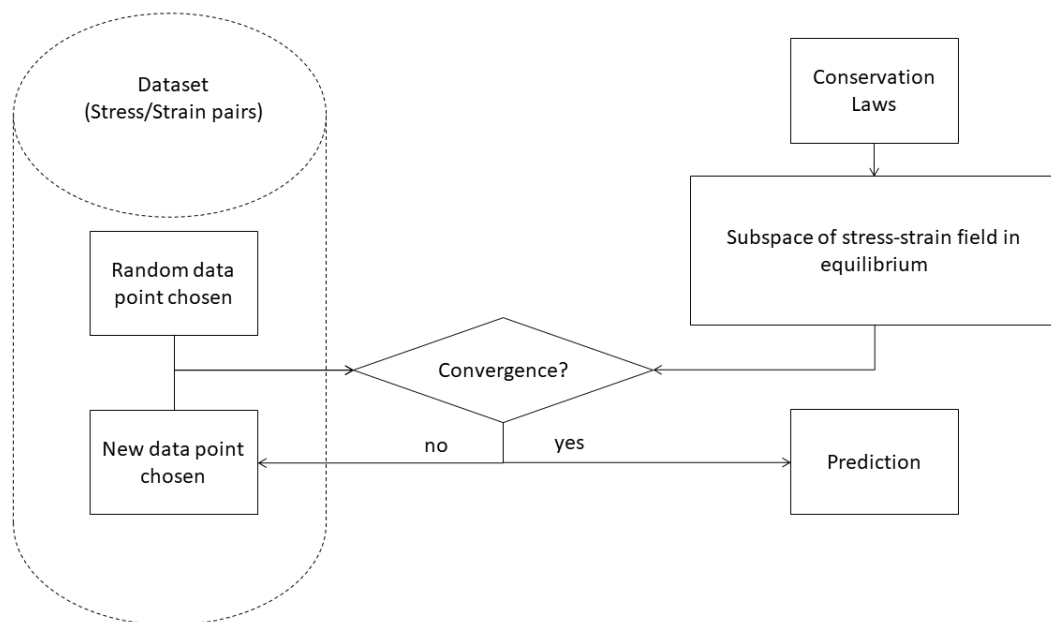
#### **4.1 Constitutive modelling with purely data-driven methods (model-free computing)**

With the increasing amount of available material data, the use of purely data-driven methods to describe materials seems more feasible. Kirchdoerfer and Ortiz (Kirchdoerfer and Ortiz, 2016) argue that these methods have the potential to represent material properties more accurately, since whenever a model is created, some information present in the data is lost and uncertainty is added to the solution. Consequently, they introduce a novel concept known as "data-driven computing." This approach relies solely on the existing material data, represented as strain-stress couples referred to as "states," and adheres to fundamental constraints and conservation laws. This approach consists of assigning two states to an element, called mechanical state and material state. A mechanical state fully respects the problem constraints and conservation laws, whereas a material state is a point from the material dataset. The solver tries to find the pair of mechanical and material states that are closest to each other, as shown in Figure 9. The application of this methodology was exemplified in two problems, static equilibrium of a nonlinear three-dimensional truss structure and linear elasticity. In both cases, the results were satisfactory. The data-driven computing approach has been extended to dynamics (Kirchdoerfer and Ortiz, 2018), by incorporating time integration, to fracture (Carrara et al., 2020) and to inelasticity problems (Eggersmann et al., 2019), which required the use of history dependent material data. Multiple strategies to represent this history dependence were considered, and the method was applied to represent both viscoelasticity and plasticity, once again in a truss structure problem. For plasticity, linear hardening was considered. The results show that the amount of data required to obtain satisfactory results significantly increases with the complexity of the problem. Other examples of the application of this methodology include the works of Karapiperis et al. and Leygue et al. (Karapiperis et al., 2021; Leygue et al., 2018). This model-free approach provides an alternative to the more widely used approaches, as presented in the previous sections. The main argument for implementing the model-free approach comes from the increased flexibility it provides, and its ability to retain all the information present in the available data, avoiding the loss of information inherent to the creation of a model.

So far, the problems used to test its application have been relatively simple (e.g., truss structures). Also, the problem of data quantity and quality dependency mentioned for the constitutive model replacement strategy, can be found for this methodology as well, with more prominence. With the lack of a pre-established metamodel, the effects of noise in the data can be even more noticeable and are more likely to lead to instabilities during numerical

simulations. Therefore, the amount and quality of data required for the successful application of this methodology to complex problems may be a prohibitive factor. Consequently, in most cases, this methodology is still not a viable alternative to the use of phenomenological constitutive models.

Another promising approach to model-free constitutive modelling is presented by Flaschel et al. (Flaschel et al., 2023, 2022), in the form of an unsupervised strategy based on full-fields data, which does not require any stress data. This strategy has been successfully applied for various elastoplastic materials and provides guarantees of thermodynamic consistency.



**Fig.9** Schematic representation of the decision algorithm proposed in (Kirchdoerfer and Ortiz, 2016).

## 4.2 Physics-informed machine learning

Another avenue for future research comes from the application of physics knowledge during the training of the metamodels. The so-called physics-based, or physics-informed machine learning is already an area of interest, as it may be a solution for some of the challenges identified in the works presented previously, such as model interpretability and guarantee of yield surface convexity (Fuhg et al., 2024). He et al. (He et al., 2021) presents a physics-constrained methodology that combines neural networks with a data-driven model free approach. This approach was applied to a beam problem, and then to the modelling of a biological material. Fuhg et al. (Fuhg et al., 2024) proposed a methodology focused on pruning

physics-informed neural network constitutive models, in an effort to reduce the number of network parameters, and as a result, increase model interpretability. This approach was applied successfully to hyperelasticity and elastoplasticity.

### **4.3 Hybrid constitutive modelling**

The concept of hybrid constitutive modelling consists of a two-step approach, in which a traditional phenomenological constitutive model is corrected, locally, by a ML metamodel. Fuhg et al. (Fuhg et al., 2023) present an application of this technique, considering multiple ML algorithms as correctors (MLP, GP and SVM). This work reproduced anisotropic yield surfaces, based on numerical data, successfully. The authors conclude that this approach can offer great versatility when solving constitutive modelling problems. Ghnatios et al. (Ghnatios et al., 2024) developed a methodology to correct a yield surface, that is based on von Mises, considering that the real surface is described by the CPB06 criterion. The methodology was applied to titanium grades with varying levels of anisotropy. Other examples of this methodology include the works of Ibáñez et al. (Ibáñez et al., 2019) and Frankel et al. (Frankel et al., 2022). The first work includes a successful approximation of a Barlat Yld04 yield function, by correction of a Hill function, while the second work focusses on the constitutive modelling of homogenized foams. In this case, the methodology was validated for multiple strain paths (uniaxial strain, simple shear and compression shear). Of particular interest is the work of Champaney et al. (Champaney et al., 2022), which gives a brief overview of this methodology, and addresses the possibilities for inclusion of physics-informed machine learning approaches.

### **4.4 Microstructural constitutive modelling**

Finally, the applications presented in this work focus primarily on the description of the plastic behaviour of materials, through phenomenological constitutive models. Other areas of constitutive modelling, such as material microstructure, can also benefit from the techniques presented, and in this context, some recent works are emerging. The inclusion of microstructural information in the constitutive modelling of metal sheets can be of particular interest, given the grain structure typical of such materials, allowing the constitutive models to better represent material anisotropy.

Schmidt and Hartmaier (Schmidt and Hartmaier, 2023) present a ML based description of crystallographic texture, which identify the relationship between material texture and the anisotropic coefficients of the yield criterion (in this case, Yld2004-18p (Barlat et al.,

2005)). Later, Shoghi and Hartmaier, (Shoghi and Hartmaier, 2024) successfully developed a model of polycrystalline material behaviour, which incorporated strain behaviour, based on the application of Support Vector Classification algorithm. To achieve this, a numerical dataset was constructed with data from multiple loading directions, applied to a Representative Volume Element (RVE) with random crystallographic texture. The resulting metamodel was validated for multiple stress states on the  $\pi$ -plane, considering various level of accumulated plastic strain.

The work of Fugh et al. (Fuhg et al., 2021) is also particularly interesting, combining microstructural information with the hybrid constitutive model approach presented previously. The microstructural data, obtained from numerical models of RVEs, considering hyperelastic materials, is used to correct the macroscale constitutive model. Other examples of ML application to constitutive modelling with microstructural information include the works of Yamanaka et al. (Yamanaka et al., 2020), Liu et al. (Liu et al., 2023) Heidenreich et al. (Heidenreich et al., 2023), Bonatti et al. (Bonatti et al., 2022) and Nascimento et al. (Nascimento et al., 2023).

## **5. Conclusion**

This work presents a summary of the state of the art regarding the application of Machine Learning to constitutive modelling for sheet metals. Two main methodologies were identified and explored. The first methodology consists of applying the ML algorithms to the identification of the constitutive model parameters for a given material. This methodology is mostly limited by the versatility of the chosen constitutive model.

The second methodology focuses on using a ML generated metamodel to replace the traditional constitutive model. These metamodels offer more flexibility in terms of the level of complexity of the material behaviour that they can describe, when compared to the traditional constitutive models. However, a drawback of this approach is that the data used to train the metamodels is material specific, making it difficult to create models applicable to different materials.

Besides the two methodologies explored, many different avenues of possible future work are presented, focusing on ML applications that have already shown promise in various adjacent research fields, namely model-free approaches, physics-informed machine learning, hybrid constitutive modelling and microstructure informed constitutive models. These approaches should be considered for future applications in the area of ML assisted sheet metal constitutive modelling.

Author	Mechanical test / Forming process	Inputs measured	Constitutive model(s)	ML algorithm(s)	Materials	Problem outputs	Performance metrics	Numerical/ experimental results
Cruz et al., 2021	Three-point bending test	Punch force- displacement curves	Swift Hardening Law	ANN	DP590 steel	K and n parameters from Swift hardening law	RMSE (root mean square error)	Numerical
Huang et al., 2017	Uniaxial tensile test	Force- displacement curves	Swift Hardening Law + Lemaitre damage model	GP (Kriging)	Aluminium AA2017 T4	K, $\varepsilon_0$ and n parameters from Swift hardening law + $\varepsilon_D^p$ , r and $D_C$ parameters from the damage model	MSE, RMSE and $R^2$	Numerical
Aguir et al., 2011	Uniaxial tensile test Plane tensile test + Circular bulge test and elliptical bulge test	Displacements	Karafillis and Boyce Criterion + Voce Hardening law	MLP + Genetic Algorithm	AISI 304 stainless steel	$R_{sat}$ , $C_R$ and $\sigma_0$ parameters from the Voce hardening law; $r_0$ , $r_{45}$ and $r_{90}$ coefficients; m and c parameters from Karafillis and Boyce criterion	None (qualitative comparison of load/displacement and pressure/height curves)	Both
Yao et al., 2021	Hot uniaxial tensile test	Stress/strain curves	Viscoplastic + Damage model	MLP + Genetic Algorithm	AA6061 aluminium	11 Viscoplastic model parameters and 5 Damage model parameters	Correlation coefficient R and RMSE	Experimental

Guo et al., 2021	Uniaxial tensile test	Strain fields	Hooke's law + Exponential hardening law	CNN + LSTM	AA6061 aluminium (for testing)	E and $\nu$ parameters from Hooke's law + 3 hardening parameters	MSRE (mean square relative error)	Both
Morand and Helm, 2019	Uniaxial tensile test	Stress/strain curves	Multiple term exponential hardening law	Single MLP and MLP ensemble	AA6014 aluminium, DC04 steel and DP600 steel	Young's Modulus, initial yield strength and hardening law parameters (4 in this work, considering that the law has two terms)	MAE (mean absolute error)	Both

Table 2. Summary of the ML applications to constitutive model parameter identification discussed.

Author	Mechanical test / Forming process	Constitutive model(s)	ML algorithm(s)	Materials	Performance metrics	Numerical/ experimental results
Zhang and Mohr, 2020	Notched tension test	Hooke's law + von Mises + Swift Hardening Law	MLP	Isotropic material (with Swift hardening law parameters mimicking a DP steel)	MSE	Numerical
Jang et al., 2021	Dog-bone tensile test; circular cup drawing	Hooke's law + von Mises + Swift Hardening Law	MLP	Isotropic material	MSE	Numerical
Hartmaier, 2020	Simple loading cases (uniaxial stress in two directions, bi-axial strain under plain stress and pure shear) applied to simple, four element numerical model	Hooke's law + Hill-like criterion (based only on principal stresses) / Tresca yield function	SVM	Steel (elastic parameters match)	$R^2$ , relative error	Numerical
Pandya et al., 2020	Uniaxial tension test, notched specimen tension test, central hole tension test and in-plane shear test	Modified Johnson-Cook model (with Swift-Voce hardening law) + von Mises	MLP	AA7075 aluminium	None (qualitative comparisons of results)	Both
Park et al., 2021	Warm V-bending	None	MLP (RBF activation function)	DP 980 steel	RMSE	Both

Gorji et al., 2020	Uniaxial tension test + 3D foam cell tested for multiple loading paths	Yld2000 criterion + HAH hardening	MLP + RNN	AA5182 aluminium + DC05 steel	MSE	Numerical
Fazily and Yoon, 2023	Uniaxial tensile and shear tests of single element blank + Cylindrical cup drawing	Hill'48 criterion / Yld2000 criterion + Swift hardening law	MLP	AA6022 aluminium + Type 409 stainless steel	MSE + MAE	Numerical
Li et al., 2023	Uniaxial tensile test	Exponential hardening law / Johnson-Cook model	Encoder: RNN, CNN, Transformer Decoder: MLP	Aluminium + AISI 316L steel	RMSE	Both
Pham et al., 2022	Uniaxial tensile test + V-shape test + truncated cone test	Hardening law proposed by Pham and Kim (Pham and Kim, 2017) + von Mises	MLP	AA5052 aluminium	MSE + $R^2$	Both
Kablman et al., 2021	Hot compression test	MD <sup>2</sup> M model	SR	AA6082 aluminium	MAE + RMSE	Both
Heidenreich and Mohr, 2024	Circular cup drawing	von Mises + Swift/Voce hardening law	RNN	Isotropic material	MSE + Median absolute error	Numerical

Table 3. Summary of the various works focused on the replacement of constitutive models with ML based metamodels.



## Symbology

Symbol	Description
$f(\mathbf{x})$	GP random variable
$\bar{\mathbf{f}}_*, \mathbf{cov}(\mathbf{f}_*)$	Vector of GP predictions, covariance of GP outputs
$i, j$	Index variables
$\mathbf{I}$	Identity matrix
$\mathbf{K}$	Covariance matrix
$V$	SVM trade-off variable
$\mathbf{w}$	Normal weight vector
$w, d$	Weight parameter, bias
$\mathbf{X}, \mathbf{X}_*$	Training matrix, testing matrix
$\mathbf{x}^t, \mathbf{x}^p$	Vector of input variables of the training and testing points, respectively
$y(\mathbf{x})$	Observed outputs for set of inputs $\mathbf{x}$
$z, z'$	Output of the current MLP node, output of a node at the previous layer
$\beta_0$	Linear function coefficient
$\gamma$	SVM threshold parameter
$\epsilon$	Noise variable
$\xi_i, \xi_i^*$	SVM slack variables
$\sigma_\epsilon^2$	Noise variance
$\emptyset$	Activation function in MLP

## Acronyms

Symbol	Description
ANN	Artificial Neural Network
CEGM	Constitutive Equation Gap Method
CNN	Convolutional Neural Network
CPB'06	Cazacu-Plunkett-Barlat 2006 yield criterion
DIC	Digital Image Correlation
DT	Decision Tree
FEM	Finite Element Method
FEMU	Finite Element Model Updating

GBDT	Gradient Boosting Decision Trees
GP	Gaussian Process
LSTM	Long Short-Term Memory
MAE	Mean Absolute Error
ML	Machine Learning
MLP	Multi-Layer Perceptron
MSE	Mean Square Error
MSRE	Mean Square Relative Error
RF	Random Forest
RMSE	Root Mean Square Error
RNN	Recurrent Neural Network
RVE	Representative Volume Element
SR	Symbolic Regression
SVM	Support Vector Machines
VFM	Virtual Fields Method

## Acknowledgements

The authors acknowledge the financial support of the Research Fund for Coal and Steel under grant agreement No 888153, by FEDER funds through the program COMPETE (Programa Operacional Factores de Competitividade), the Portuguese Foundation for Science and Technology (FCT) under projects UIDB/00285/2020, UIDB/00326/2020, UIDB/00481/2020, UIDP/00481/2020-FCT DOI 10.54499/UIDB/00481/2020 (<https://doi.org/10.54499/UIDB/00481/2020>) and DOI 10.54499/UIDP/00481/2020 (<https://doi.org/10.54499/UIDP/00481/2020>), CENTRO-01-0145-FEDER-022083, LA/P/0104/2020, LA/P/0112/2020, project RealForm (2022.02370.PTDC) and 2022.05783.PTDC. The first author, Armando E. Marques, was supported by a grant for scientific research from the Portuguese Foundation for Science and Technology (<https://doi.org/10.54499/2020.08449.BD>). All supports are gratefully acknowledged.

## Author Contributions

**Armando E. Marques:** Conceptualization, Investigation, Writing – Original Draft, Writing – Review & Editing. **Tomás G. Parreira:** Investigation, Writing – Review & Editing. **André F. G. Pereira:** Writing – Review & Editing, Supervision. **Bernardete M. Ribeiro:** Supervision, Funding Acquisition. **Pedro A. Prates:** Conceptualization, Writing – Review & Editing, Supervision, Funding Acquisition.

## Declaration of competing interests

The authors declare no conflict of interests.

## Data availability

No data was used for the research in this work.

## References

- Aguir, H., BelHadjSalah, H., Hambli, R., 2011. Parameter identification of an elasto-plastic behaviour using artificial neural networks-genetic algorithm method. *Materials and Design* 32, 48–53. <https://doi.org/10.1016/j.matdes.2010.06.039>
- Albawi, S., Mohammed, T.A., Al-Zawi, S., 2018. Understanding of a convolutional neural network. *Proceedings of 2017 International Conference on Engineering and Technology, ICET 2017 2018-Janua*, 1–6. <https://doi.org/10.1109/ICEngTechnol.2017.8308186>
- Aljuaid, A., Anwar, M., 2022. Survey of Supervised Learning for Medical Image Processing. *SN Computer Science* 3, 1–22. <https://doi.org/10.1007/s42979-022-01166-1>
- Angelis, D., Sofos, F., Karakasidis, T.E., 2023. Artificial Intelligence in Physical Sciences: Symbolic Regression Trends and Perspectives. *Archives of Computational Methods in Engineering* 30, 3845–3865. <https://doi.org/10.1007/s11831-023-09922-z>
- Asadzadeh, M.Z., Gänser, H.P., Mücke, M., 2021. Symbolic regression based hybrid semiparametric modelling of processes: An example case of a bending process.

- Applications in Engineering Science 6, 100049.  
<https://doi.org/10.1016/j.apples.2021.100049>
- Augusto, D.A., Barbosa, H.J.C., 2000. Symbolic regression via genetic programming. Proceedings - Brazilian Symposium on Neural Networks, SBRN 173–178.  
<https://doi.org/10.1109/SBRN.2000.889734>
- Avril, S., Bonnet, M., Bretelle, A.S., Grédiac, M., Hild, F., Ienny, P., Latourte, F., Lemosse, D., Pagano, S., Pagnacco, E., Pierron, F., 2008. Overview of identification methods of mechanical parameters based on full-field measurements. *Experimental Mechanics* 48, 381–402. <https://doi.org/10.1007/s11340-008-9148-y>
- Banabic, D., Barlat, F., Cazacu, O., Kuwabara, T., 2020. Advances in anisotropy of plastic behaviour and formability of sheet metals. *International Journal of Material Forming* 13, 749–787. <https://doi.org/10.1007/s12289-020-01580-x>
- Barlat, F., Aretz, H., Yoon, J.W., Karabin, M.E., Brem, J.C., Dick, R.E., 2005. Linear transformation-based anisotropic yield functions. *International Journal of Plasticity* 21, 1009–1039. <https://doi.org/10.1016/j.ijplas.2004.06.004>
- Barlat, F., Brem, J.C., Yoon, J.W., Chung, K., Dick, R.E., Lege, D.J., Pourboghrat, F., Choi, S.H., Chu, E., 2003. Plane stress yield function for aluminum alloy sheets - Part 1: Theory. *International Journal of Plasticity* 19, 1297–1319.  
[https://doi.org/10.1016/S0749-6419\(02\)00019-0](https://doi.org/10.1016/S0749-6419(02)00019-0)
- Bomarito, G.F., Townsend, T.S., Stewart, K.M., Esham, K. V., Emery, J.M., Hochhalter, J.D., 2021. Development of interpretable, data-driven plasticity models with symbolic regression. *Computers and Structures* 252, 106557.  
<https://doi.org/10.1016/j.compstruc.2021.106557>
- Bonatti, C., Berisha, B., Mohr, D., 2022. From CP-FFT to CP-RNN: Recurrent neural network surrogate model of crystal plasticity. *International Journal of Plasticity* 158, 103430. <https://doi.org/10.1016/j.ijplas.2022.103430>
- Bouchard, P.O., Bourgeon, L., Fayolle, S., Mocellin, K., 2011. An enhanced Lemaitre model formulation for materials processing damage computation. *International Journal of Material Forming* 4, 299–315. <https://doi.org/10.1007/s12289-010-0996-5>
- Breiman L., 2001. Random Forests. *Machine Learning* 45, 5–32.  
<https://doi.org/doi.org/10.1023/A:1010933404324>
- Brito, L.C., Susto, G.A., Brito, J.N., Duarte, M.A.V., 2021. Fault detection of bearing: An unsupervised machine learning approach exploiting feature extraction and dimensionality reduction. *Informatics* 8. <https://doi.org/10.3390/informatics8040085>

- Brodnik, N.R., Muir, C., Tulshibagwale, N., Rossin, J., Echlin, M.P., Hamel, C.M., Kramer, S.L.B., Pollock, T.M., Kiser, J.D., Smith, C., Daly, S.H., 2023. Perspective: Machine learning in experimental solid mechanics. *Journal of the Mechanics and Physics of Solids* 173, 105231. <https://doi.org/10.1016/j.jmps.2023.105231>
- Budhiraja, S.S., Mago, V., 2020. A supervised learning approach for heading detection. *Expert Systems* 37, 1–15. <https://doi.org/10.1111/exsy.12520>
- Buhuş, E.R., Grama, L., Şerbu, C., 2017. A facial recognition application based on incremental supervised learning. *Proceedings - 2017 IEEE 13th International Conference on Intelligent Computer Communication and Processing, ICCP 2017* 279–286. <https://doi.org/10.1109/ICCP.2017.8117017>
- Carrara, P., De Lorenzis, L., Stainier, L., Ortiz, M., 2020. Data-driven fracture mechanics. *Computer Methods in Applied Mechanics and Engineering* 372, 113390. <https://doi.org/10.1016/j.cma.2020.113390>
- Cazacu, O., Barlat, F., 2001. Generalization of Drucker's yield criterion to orthotropy. *Mathematics and Mechanics of Solids* 6, 613–630. <https://doi.org/10.1177/108128650100600603>
- Cazacu, O., Plunkett, B., Barlat, F., 2006. Orthotropic yield criterion for hexagonal closed packed metals. *International Journal of Plasticity* 22, 1171–1194. <https://doi.org/10.1016/j.ijplas.2005.06.001>
- Champaney, V., Chinesta, F., Cueto, E., 2022. Engineering empowered by physics-based and data-driven hybrid models: A methodological overview. *International Journal of Material Forming* 15, 1–14. <https://doi.org/10.1007/s12289-022-01678-4>
- Chaparro, B.M., Thuillier, S., Menezes, L.F., Manach, P.Y., Fernandes, J. V., 2008. Material parameters identification: Gradient-based, genetic and hybrid optimization algorithms. *Computational Materials Science* 44, 339–346. <https://doi.org/10.1016/j.commatsci.2008.03.028>
- Cruz, D.J., Barbosa, M.R., Santos, A.D., Miranda, S.S., Amaral, R.L., 2021. Application of Machine Learning to Bending Processes and Material Identification. *Metals* 11, 1418. <https://doi.org/10.3390/met11091418>
- D'Angelo, G., Palmieri, F., 2021. GGA: A modified genetic algorithm with gradient-based local search for solving constrained optimization problems. *Information Sciences* 547, 136–162. <https://doi.org/10.1016/j.ins.2020.08.040>
- Diller, M., Thomas, W., Ahmetoglu, M.A., Akgerman, N., Altan, T., 1997. Applications of computer simulations for part and process design for automotive stampings. *SAE*

- Technical Papers. <https://doi.org/10.4271/970985>
- Dornheim, J., Morand, L., Nallani, H.J., Helm, D., 2024. Neural Networks for Constitutive Modeling: From Universal Function Approximators to Advanced Models and the Integration of Physics, *Archives of Computational Methods in Engineering*. Springer Netherlands. <https://doi.org/10.1007/s11831-023-10009-y>
- Dreyfus, S.E., 1990. Artificial Neural Networks, Back Propagation, and the Kelley-Bryson Gradient Procedure. *Journal of Guidance, Control, and Dynamics* 13, 926–928. <https://doi.org/10.2514/3.25422>
- Eggersmann, R., Kirchdoerfer, T., Reese, S., Stainier, L., Ortiz, M., 2019. Model-Free Data-Driven inelasticity. *Computer Methods in Applied Mechanics and Engineering* 350, 81–99. <https://doi.org/10.1016/j.cma.2019.02.016>
- Fazily, P., Yoon, J.W., 2023. Machine learning-driven stress integration method for anisotropic plasticity in sheet metal forming. *International Journal of Plasticity* 166, 103642. <https://doi.org/10.1016/j.ijplas.2023.103642>
- Feng, Y., Hong, Z., Gao, Y., Lu, R., Wang, Y., Tan, J., 2019. Optimization of variable blank holder force in deep drawing based on support vector regression model and trust region. *International Journal of Advanced Manufacturing Technology* 105, 4265–4278. <https://doi.org/10.1007/s00170-019-04477-5>
- Flaschel, M., Kumar, S., De Lorenzis, L., 2023. Automated discovery of generalized standard material models with EUCLID. *Computer Methods in Applied Mechanics and Engineering* 405, 115867. <https://doi.org/10.1016/j.cma.2022.115867>
- Flaschel, M., Kumar, S., De Lorenzis, L., 2022. Discovering plasticity models without stress data. *npj Computational Materials* 8, 1–10. <https://doi.org/10.1038/s41524-022-00752-4>
- Frankel, A., Hamel, C.M., Bolintineanu, D., Long, K., Kramer, S., 2022. Machine learning constitutive models of elastomeric foams. *Computer Methods in Applied Mechanics and Engineering* 391, 114492. <https://doi.org/10.1016/j.cma.2021.114492>
- Friedman, J.H., 2001. Greedy function approximation: A gradient boosting machine. *Annals of Statistics* 29, 1189–1232. <https://doi.org/10.1214/aos/1013203451>
- Fu, J., Barlat, F., Kim, J.H., Pierron, F., 2017. Application of the virtual fields method to the identification of the homogeneous anisotropic hardening parameters for advanced high strength steels. *International Journal of Plasticity* 93, 229–250. <https://doi.org/10.1016/j.ijplas.2016.07.013>
- Fuhg, J.N., Böhm, C., Bouklas, N., Fau, A., Wriggers, P., Marino, M., 2021. Model-data-driven constitutive responses: Application to a multiscale computational framework.

- International Journal of Engineering Science 167, 103522.  
<https://doi.org/10.1016/j.ijengsci.2021.103522>
- Fuhg, J.N., Fau, A., Bouklas, N., Marino, M., 2023. Enhancing phenomenological yield functions with data: Challenges and opportunities. *European Journal of Mechanics, A/Solids* 99, 104925. <https://doi.org/10.1016/j.euromechsol.2023.104925>
- Fuhg, J.N., Jones, R.E., Bouklas, N., 2024. Extreme sparsification of physics-augmented neural networks for interpretable model discovery in mechanics. *Computer Methods in Applied Mechanics and Engineering* 426. <https://doi.org/10.1016/j.cma.2024.116973>
- G.R.Johnson, W.H.Cook, 1983. A constitutive model and data for metals subjected to large strains, high strain rates and high",. *Proceedings of the 7th International Symposium on Ballistics Committee* 21, 541–547.
- Generale, A.P., Robertson, A.E., Kelly, C., Kalidindi, S.R., 2024. Inverse stochastic microstructure design. *Acta Materialia* 271, 119877.  
<https://doi.org/10.1016/j.actamat.2024.119877>
- Ghahramani, Z., 2004. LNAI 3176 - Unsupervised Learning. *Machine Learning* 72–112.
- Ghnatios, C., Cazacu, O., Revil-Baudard, B., Chinesta, F., 2024. A new methodology for anisotropic yield surface description using model order reduction techniques and invariant neural network. *Journal of the Mechanics and Physics of Solids* 184, 105542.  
<https://doi.org/10.1016/j.jmps.2024.105542>
- Gorji, M.B., Mozaffar, M., Heidenreich, J.N., Cao, J., Mohr, D., 2020. On the potential of recurrent neural networks for modeling path dependent plasticity. *Journal of the Mechanics and Physics of Solids* 143, 103972.  
<https://doi.org/10.1016/j.jmps.2020.103972>
- Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., Liu, T., Wang, X., Wang, G., Cai, J., Chen, T., 2018. Recent advances in convolutional neural networks. *Pattern Recognition* 77, 354–377. <https://doi.org/10.1016/j.patcog.2017.10.013>
- Guo, Z., Bai, R., Lei, Z., Jiang, H., Liu, D., Zou, J., Yan, C., 2021. CPINet: Parameter identification of path-dependent constitutive model with automatic denoising based on CNN-LSTM. *European Journal of Mechanics, A/Solids* 90, 104327.  
<https://doi.org/10.1016/j.euromechsol.2021.104327>
- Habraken, A.M., Aksen, T.A., Alves, J.L., Amaral, R.L., Betaieb, E., Chandola, N., Corallo, L., Cruz, D.J., Duchêne, L., Engel, B., Esener, E., Firat, M., Frohn-Sörensen, P., Galán-López, J., Ghiabakloo, H., Kestens, L.A.I., Lian, J., Lingam, R., Liu, W., Ma, J., Menezes, L.F., Nguyen-Minh, T., Miranda, S.S., Neto, D.M., Pereira, A.F.G., Prates,

- P.A., Reuter, J., Revil-Baudard, B., Rojas-Ulloa, C., Sener, B., Shen, F., Van Bael, A., Verleysen, P., Barlat, F., Cazacu, O., Kuwabara, T., Lopes, A., Oliveira, M.C., Santos, A.D., Vincze, G., 2022. Analysis of ESAFORM 2021 cup drawing benchmark of an Al alloy, critical factors for accuracy and efficiency of FE simulations. *International Journal of Material Forming* 15, 61. <https://doi.org/10.1007/s12289-022-01672-w>
- Hamel, C.M., Long, K.N., Kramer, S.L.B., 2023. Calibrating constitutive models with full-field data via physics informed neural networks. *Strain* 59, 1–31. <https://doi.org/10.1111/str.12431>
- Hartmaier, A., 2020. Data-oriented constitutive modeling of plasticity in metals. *Materials* 13, 1600. <https://doi.org/10.3390/ma13071600>
- He, X., He, Q., Chen, J.S., 2021. Deep autoencoders for physics-constrained data-driven nonlinear materials modeling. *Computer Methods in Applied Mechanics and Engineering* 385, 114034. <https://doi.org/10.1016/j.cma.2021.114034>
- Heidenreich, J.N., Gorji, M.B., Mohr, D., 2023. Modeling structure-property relationships with convolutional neural networks: Yield surface prediction based on microstructure images. *International Journal of Plasticity* 163, 103506. <https://doi.org/10.1016/j.ijplas.2022.103506>
- Heidenreich, J.N., Mohr, D., 2024. Recurrent neural network plasticity models: Unveiling their common core through multi-task learning. *Computer Methods in Applied Mechanics and Engineering* 426, 116991. <https://doi.org/10.1016/j.cma.2024.116991>
- Hill, R., 1948. A theory of the yielding and plastic flow of anisotropic metals. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* 193, 281–297. <https://doi.org/10.1098/rspa.1948.0045>
- Hou, Y., Myung, D., Park, J.K., Min, J., Lee, H.R., El-Aty, A.A., Lee, M.G., 2023. A Review of Characterization and Modelling Approaches for Sheet Metal Forming of Lightweight Metallic Materials. *Materials* 16. <https://doi.org/10.3390/ma16020836>
- Huang, C., El Hami, A., Radi, B., 2017. Metamodel-based inverse method for parameter identification: elastic–plastic damage model. *Engineering Optimization* 49, 633–653. <https://doi.org/10.1080/0305215X.2016.1206537>
- Huber, N., Tsakmakis, C., 1999a. Determination of constitutive properties from spherical indentation data using neural networks. Part II: plasticity with nonlinear isotropic and kinematic hardening. *Journal of the Mechanics and Physics of Solids* 47, 1589–1607. [https://doi.org/10.1016/S0022-5096\(98\)00109-4](https://doi.org/10.1016/S0022-5096(98)00109-4)
- Huber, N., Tsakmakis, C., 1999b. Determination of constitutive properties from spherical



- indentation data using neural networks. Part I: The case of pure kinematic hardening in plasticity laws. *Journal of the Mechanics and Physics of Solids* 47, 1569–1588.  
[https://doi.org/10.1016/S0022-5096\(98\)00109-4](https://doi.org/10.1016/S0022-5096(98)00109-4)
- Ibáñez, R., Abisset-Chavanne, E., González, D., Duval, J.L., Cueto, E., Chinesta, F., 2019. Hybrid constitutive modeling: data-driven learning of corrections to plasticity models. *International Journal of Material Forming* 12, 717–725. <https://doi.org/10.1007/s12289-018-1448-x>
- Jamli, M.R., Ariffin, A.K., Wahab, D.A., 2015. Incorporating feedforward neural network within finite element analysis for L-bending springback prediction. *Expert Systems with Applications* 42, 2604–2614. <https://doi.org/10.1016/j.eswa.2014.11.005>
- Jang, D.P., Fazily, P., Yoon, J.W., 2021. Machine learning-based constitutive model for J2-plasticity. *International Journal of Plasticity* 138, 102919.  
<https://doi.org/10.1016/j.ijplas.2020.102919>
- Jin, H., Zhang, E., Espinosa, H.D., 2023. Recent Advances and Applications of Machine Learning in Experimental Solid Mechanics: A Review. *Applied Mechanics Reviews* 75.  
<https://doi.org/10.1115/1.4062966>
- Jingdong, L., Li, H., Hongbo, Z., 2017. Forming defects prediction for sheet metal forming using Gaussian process regression. *Proceedings of the 29th Chinese Control and Decision Conference, CCDC 2017* 472–476.  
<https://doi.org/10.1109/CCDC.2017.7978140>
- Kablman, E., Kolody, A.H., Kronsteiner, J., Kommenda, M., Kronberger, G., 2021. Application of symbolic regression for constitutive modeling of plastic deformation. *Applications in Engineering Science* 6, 100052.  
<https://doi.org/10.1016/j.apples.2021.100052>
- Karafillis, A.P., Boyce, M.C., 1993. A general anisotropic yield criterion using bounds and a transformation weighting tensor. *Journal of the Mechanics and Physics of Solids* 41, 1859–1886. [https://doi.org/10.1016/0022-5096\(93\)90073-O](https://doi.org/10.1016/0022-5096(93)90073-O)
- Karapiperis, K., Stainier, L., Ortiz, M., Andrade, J.E., 2021. Data-Driven multiscale modeling in mechanics. *Journal of the Mechanics and Physics of Solids* 147, 104239.  
<https://doi.org/10.1016/j.jmps.2020.104239>
- Kirchdoerfer, T., Ortiz, M., 2018. Data-driven computing in dynamics. *International Journal for Numerical Methods in Engineering* 113, 1697–1710.  
<https://doi.org/10.1002/nme.5716>
- Kirchdoerfer, T., Ortiz, M., 2016. Data-driven computational mechanics. *Computer Methods*

- in *Applied Mechanics and Engineering* 304, 81–101.  
<https://doi.org/10.1016/j.cma.2016.02.001>
- Kubik, C., Knauer, S.M., Groche, P., 2022. Smart sheet metal forming: importance of data acquisition, preprocessing and transformation on the performance of a multiclass support vector machine for predicting wear states during blanking. *Journal of Intelligent Manufacturing* 33, 259–282. <https://doi.org/10.1007/s10845-021-01789-w>
- Kunka, C., Shanker, A., Chen, E.Y., Kalidindi, S.R., Dingreville, R., 2021. Decoding defect statistics from diffractograms via machine learning. *npj Computational Materials* 7. <https://doi.org/10.1038/s41524-021-00539-z>
- Latourte, F., Chrysochoos, A., Pagano, S., Wattrisse, B., 2008. Elastoplastic behavior identification for heterogeneous loadings and materials. *Experimental Mechanics* 48, 435–449. <https://doi.org/10.1007/s11340-007-9088-y>
- Lemaitre, J., Chaboche, J.-L., 1990. Plasticity, in: *Mechanics of Solid Materials*. Cambridge University Press, pp. 161–252. <https://doi.org/10.1017/CBO9781139167970>
- Leygue, A., Coret, M., Réthoré, J., Stainier, L., Verron, E., 2018. Data-based derivation of material response. *Computer Methods in Applied Mechanics and Engineering* 331, 184–196. <https://doi.org/10.1016/j.cma.2017.11.013>
- Li, D. wei, Liu, J. xiang, Fan, Y. sheng, Yang, X. guang, Huang, W. qing, 2024. A preliminary discussion about the application of machine learning in the field of constitutive modeling focusing on alloys. *Journal of Alloys and Compounds* 976, 173210. <https://doi.org/10.1016/j.jallcom.2023.173210>
- Li, Q.J., Cinbiz, M.N., Zhang, Y., He, Q., Beausoleil, G., Li, J., 2023. Robust deep learning framework for constitutive relations modeling. *Acta Materialia* 254, 118959. <https://doi.org/10.1016/j.actamat.2023.118959>
- Li, X., Roth, C.C., Mohr, D., 2023. Neural network based rate- and temperature-dependent Hosford–Coulomb fracture initiation model. *International Journal of Mechanical Sciences* 260, 108643. <https://doi.org/10.1016/j.ijmecsci.2023.108643>
- Liu, B., Ocegueda, E., Trautner, M., Stuart, A.M., Bhattacharya, K., 2023. Learning macroscopic internal variables and history dependence from microscopic models. *Journal of the Mechanics and Physics of Solids* 178, 1–34. <https://doi.org/10.1016/j.jmps.2023.105329>
- Liu, B., Zhang, Q., Gielen, G.G.E., 2014. A gaussian process surrogate model assisted evolutionary algorithm for medium scale expensive optimization problems. *IEEE Transactions on Evolutionary Computation* 18, 180–192.

- <https://doi.org/10.1109/TEVC.2013.2248012>
- Liu, S., Xia, Y., Shi, Z., Yu, H., Li, Z., Lin, J., 2021. Deep Learning in Sheet Metal Bending with a Novel Theory-Guided Deep Neural Network. *IEEE/CAA Journal of Automatica Sinica* 8, 565–581. <https://doi.org/10.1109/JAS.2021.1003871>
- Lourenço, R., Andrade-Campos, A., Georgieva, P., 2022. The Use of Machine-Learning Techniques in Material Constitutive Modelling for Metal Forming Processes. *Metals* 12, 427. <https://doi.org/10.3390/met12030427>
- Marques, A.E., Prates, P.A., Pereira, A.F.G., Oliveira, M.C., Fernandes, J. V., Ribeiro, B.M., 2020. Performance comparison of parametric and non-parametric regression models for uncertainty analysis of sheet metal forming processes. *Metals* 10, 457. <https://doi.org/10.3390/met10040457>
- Martín-Guerrero, J.D., Lamata, L., 2021. Reinforcement Learning and Physics. *Applied Sciences* 11, 8589. <https://doi.org/10.3390/app11188589>
- Mecking, H., Kocks, U.F., 1981. Kinetics of flow and strain-hardening. *Acta Metallurgica* 29, 1865–1875. [https://doi.org/10.1016/0001-6160\(81\)90112-7](https://doi.org/10.1016/0001-6160(81)90112-7)
- Morand, L., Helm, D., 2019. A mixture of experts approach to handle ambiguities in parameter identification problems in material modeling. *Computational Materials Science* 167, 85–91. <https://doi.org/10.1016/j.commatsci.2019.04.003>
- Müller, A., Karathanasopoulos, N., Roth, C.C., Mohr, D., 2021. Machine Learning Classifiers for Surface Crack Detection in Fracture Experiments. *International Journal of Mechanical Sciences* 209. <https://doi.org/10.1016/j.ijmecsci.2021.106698>
- Murtagh, F., 1991. Multilayer perceptrons for classification and regression. *Neurocomputing* 2, 183–197. [https://doi.org/10.1016/0925-2312\(91\)90023-5](https://doi.org/10.1016/0925-2312(91)90023-5)
- Muth, A., Venkatraman, A., John, R., Pilchak, A., Kalidindi, S.R., McDowell, D.L., 2023. Neighborhood spatial correlations and machine learning classification of fatigue hot-spots in Ti–6Al–4V. *Mechanics of Materials* 182, 104679. <https://doi.org/10.1016/j.mechmat.2023.104679>
- Nascimento, A., Roongta, S., Diehl, M., Beyerlein, I.J., 2023. A machine learning model to predict yield surfaces from crystal plasticity simulations. *International Journal of Plasticity* 161, 103507. <https://doi.org/10.1016/j.ijplas.2022.103507>
- Nath, V., Chattopadhyay, C., 2021. S2D2Net: an Improved Approach for Robust Steel Surface Defects Diagnosis With Small Sample Learning. *Proceedings - International Conference on Image Processing, ICIP 2021-Septe*, 1199–1203. <https://doi.org/10.1109/ICIP42928.2021.9506405>

- Nketiah, E.A., Chenlong, L., Yingchuan, J., Aram, S.A., 2023. Recurrent neural network modeling of multivariate time series and its application in temperature forecasting. *PLoS ONE* 18, 1–18. <https://doi.org/10.1371/journal.pone.0285713>
- Pandya, K.S., Roth, C.C., Mohr, D., 2020. Strain rate and temperature dependent fracture of aluminum alloy 7075: Experiments and neural network modeling. *International Journal of Plasticity* 135, 102788. <https://doi.org/10.1016/j.ijplas.2020.102788>
- Park, S.H., Lee, E.H., Choi, H., Lee, J.Y., 2021. A Thermal–Elastic–Plastic Constitutive Model using the Radial Basis Function Neural Network and Application for an Energy Efficient Warm Forming Process. *International Journal of Precision Engineering and Manufacturing* 22, 329–343. <https://doi.org/10.1007/s12541-020-00460-3>
- Périé, J.N., Leclerc, H., Roux, S., Hild, F., 2009. Digital image correlation and biaxial test on composite material for anisotropic damage law identification. *International Journal of Solids and Structures* 46, 2388–2396. <https://doi.org/10.1016/j.ijsolstr.2009.01.025>
- Pham, Q.T., Kim, Y.S., 2017. Identification of the plastic deformation characteristics of AL5052-O sheet based on the non-associated flow rule. *Metals and Materials International* 23, 254–263. <https://doi.org/10.1007/s12540-017-6378-5>
- Pham, Q.T., Le, H.S., Nguyen, A.T., Xiao, X., Kim, Y.S., Nguyen, V.D., Tran, H.S., Van Tran, X., 2022. A machine learning–based methodology for identification of the plastic flow in aluminum sheets during incremental sheet forming processes. *International Journal of Advanced Manufacturing Technology* 120, 3559–3584. <https://doi.org/10.1007/s00170-022-08698-z>
- Plunkett, B., Cazacu, O., Barlat, F., 2008. Orthotropic yield criteria for description of the anisotropy in tension and compression of sheet metals. *International Journal of Plasticity* 24, 847–866. <https://doi.org/10.1016/j.ijplas.2007.07.013>
- Prates, P.A., Pereira, A.F.G., Sakharova, N.A., Oliveira, M.C., Fernandes, J. V., 2016. Inverse Strategies for Identifying the Parameters of Constitutive Laws of Metal Sheets. *Advances in Materials Science and Engineering* 2016, 4152963. <https://doi.org/10.1155/2016/4152963>
- Rabahallah, M., Balan, T., Bouvier, S., Bacroix, B., Barlat, F., Chung, K., Teodosiu, C., 2009. Parameter identification of advanced plastic strain rate potentials and impact on plastic anisotropy prediction. *International Journal of Plasticity* 25, 491–512. <https://doi.org/10.1016/j.ijplas.2008.03.006>
- Rokach, L., Maimon, O., 2005. Decision Trees, in: *Data Mining and Knowledge Discovery Handbook*. Springer-Verlag, New York, pp. 165–192. <https://doi.org/10.1007/0-387->

- Rosenkranz, M., Kalina, K.A., Brummund, J., Kästner, M., 2023. A comparative study on different neural network architectures to model inelasticity. *International Journal for Numerical Methods in Engineering* 124, 4802–4840. <https://doi.org/10.1002/nme.7319>
- Rossi, M., Lattanzi, A., Morichelli, L., Martins, J.M.P., Thuillier, S., Andrade-Campos, A., Coppeliers, S., 2022. Testing methodologies for the calibration of advanced plasticity models for sheet metals: A review. *Strain* 58, 1–28. <https://doi.org/10.1111/str.12426>
- Salehinejad, H., Sankar, S., Barfett, J., Colak, E., Valaee, S., 2017. Recent Advances in Recurrent Neural Networks 1–21.
- Sangeetha, R., Kalpana, B., 2010. A Comparative Study and Choice of an Appropriate Kernel for Support Vector Machines. *Communications in Computer and Information Science* 101, 549–553. [https://doi.org/10.1007/978-3-642-15766-0\\_93](https://doi.org/10.1007/978-3-642-15766-0_93)
- Saunders, R., Rawlings, A., Birnbaum, A., Iliopoulos, A., Michopoulos, J., Lagoudas, D., Elwany, A., 2022. Additive Manufacturing Melt Pool Prediction and Classification via Multifidelity Gaussian Process Surrogates. *Integrating Materials and Manufacturing Innovation* 11, 497–515. <https://doi.org/10.1007/s40192-022-00276-1>
- Schmidt, J., Hartmaier, A., 2023. A new texture descriptor for data-driven constitutive modeling of anisotropic plasticity. *Journal of Materials Science* 58, 14029–14050. <https://doi.org/10.1007/s10853-023-08852-2>
- Schölkopf, B., 2000. The kernel trick for distances. *Advances in Neural Information Processing Systems* 301–307.
- Senn, M., Link, N., Pollak, J., Lee, J.H., 2014. Reducing the computational effort of optimal process controllers for continuous state spaces by using incremental learning and post-decision state formulations. *Journal of Process Control* 24, 133–143. <https://doi.org/10.1016/j.jprocont.2014.01.002>
- Sherstinsky, A., 2020. Fundamentals of Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) network. *Physica D: Nonlinear Phenomena* 404, 132306. <https://doi.org/10.1016/j.physd.2019.132306>
- Shoghi, R., Hartmaier, A., 2024. A machine learning constitutive model for plasticity and strain hardening of polycrystalline metals based on data from micromechanical simulations. *Machine Learning: Science and Technology* 5. <https://doi.org/10.1088/2632-2153/ad379e>
- Smola, A.J., Schölkopf, B., 2004. A tutorial on support vector regression. *Statistics and Computing* 14, 199–222. <https://doi.org/10.1023/B:STCO.0000035301.49549.88>

- Stanev, V., Vesselinov, V. V., Kusne, A.G., Antoszewski, G., Takeuchi, I., Alexandrov, B.S., 2018. Unsupervised phase mapping of X-ray diffraction data by nonnegative matrix factorization integrated with custom clustering. *npj Computational Materials* 4. <https://doi.org/10.1038/s41524-018-0099-2>
- Swift, H.W., 1952. Plastic instability under plane stress. *Journal of the Mechanics and Physics of Solids* 1, 1–18. [https://doi.org/10.1016/0022-5096\(52\)90002-1](https://doi.org/10.1016/0022-5096(52)90002-1)
- Ur, S., Baig, R., Wasif, M., 2021. Machine learning for the prediction of springback in High tensile strength steels after V-bending process using tree-based learning. *Research Square*. <https://doi.org/10.21203/rs.3.rs-795174/v1>
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł., Polosukhin, I., 2017. Attention is all you need, in: *Advances in Neural Information Processing Systems*. pp. 5999–6009.
- Villarreal, R., Vlassis, N.N., Phan, N.N., Catanach, T.A., Jones, R.E., Trask, N.A., Kramer, S.L.B., Sun, W.C., 2023. Design of experiments for the calibration of history-dependent models via deep reinforcement learning and an enhanced Kalman filter. *Computational Mechanics* 72, 95–124. <https://doi.org/10.1007/s00466-023-02335-6>
- Voce, E., 1948. The relationship between stress and strain from homogenous deformation. *Journal of the Institute for Metals* 74, 537–562.
- von Mises, R., 1913. *Mechanik der festen Körper im plastisch deformation. Zustand.* Nachr. Ges. Wiss. Gottigen 582.
- Wang, H., Chen, L., Li, E., 2018. Time dependent sheet metal forming optimization by using Gaussian process assisted firefly algorithm. *International Journal of Material Forming* 11, 279–295. <https://doi.org/10.1007/s12289-017-1352-9>
- Wang, K., Sun, W.C., Du, Q., 2021. A non-cooperative meta-modeling game for automated third-party calibrating, validating and falsifying constitutive laws with parallelized adversarial attacks. *Computer Methods in Applied Mechanics and Engineering* 373, 113514. <https://doi.org/10.1016/j.cma.2020.113514>
- Yamanaka, A., Kamijyo, R., Koenuma, K., Watanabe, I., Kuwabara, T., 2020. Deep neural network approach to estimate biaxial stress-strain curves of sheet metals. *Materials and Design* 195, 108970. <https://doi.org/10.1016/j.matdes.2020.108970>
- Yao, D., Duan, Y. chuan, Li, M. yu, Guan, Y. ping, 2021. Hybrid identification method of coupled viscoplastic-damage constitutive parameters based on BP neural network and genetic algorithm. *Engineering Fracture Mechanics* 257, 108027. <https://doi.org/10.1016/j.engfracmech.2021.108027>

- Yoshida, F., Hamasaki, H., Uemori, T., 2013. A user-friendly 3D yield function to describe anisotropy of steel sheets. *International Journal of Plasticity* 45, 119–139.  
<https://doi.org/10.1016/j.ijplas.2013.01.010>
- Yu-Lin, L., Wen-Liang, C., Li-Ping, D., Xiao-Ping, W., 2013. Response surface methodology based on support vector regression for polygon blank shape optimization design. *International Journal of Advanced Manufacturing Technology* 66, 1397–1405.  
<https://doi.org/10.1007/s00170-012-4417-4>
- Zhang, A., Mohr, D., 2020. Using neural networks to represent von Mises plasticity with isotropic hardening. *International Journal of Plasticity* 132, 102732.  
<https://doi.org/10.1016/j.ijplas.2020.102732>
- Zhang, Y., Xu, X., 2022. Predicting springback radii and angles in air bending of high-strength sheet steel through gaussian process regressions. *International Journal on Interactive Design and Manufacturing* 16, 863–870. <https://doi.org/10.1007/s12008-022-00945-7>
- Zhou, Z.H., 2018. A brief introduction to weakly supervised learning. *National Science Review* 5, 44–53. <https://doi.org/10.1093/nsr/nwx106>