Single and ensemble classifiers for defect prediction in sheet metal forming under variability

M. A. Dib[1] · N. J. Oliveira[1] · A. E. Marques[2] · M. C. Oliveira[2] · J. V. Fernandes[2] · B. M. Ribeiro[1] ·
P. A. Prates[2]

Abstract

This paper presents an approach, based on machine learning techniques, to predict the occurrence of defects in sheet metal forming processes, exposed to sources of scatter in the material properties and process parameters. An empirical analysis of performance of ML techniques is presented, considering both single learning and ensemble models. These are trained using data sets populated with numerical simulation results of two sheet metal forming processes: U-Channel and Square Cup. Data sets were built for three distinct steel sheets. A total of eleven input features, related to the mechanical properties, sheet thickness and process parameters, were considered; also, two types of defects (outputs) were analysed for each process. The sampling data were generated, assuming that the variability of each input feature is described by a normal distribution. For a given type of defect, most single classifiers show similar performances, regardless of the material. When comparing single learning and ensemble models, the latter can provide an efficient alternative. The fact that ensemble predictive models present relatively high performances, combined with the possibility of reconciling model bias and variance, offer a promising direction for its application in industrial environment.

Keywords Machine learning · Ensemble learning · Defect prediction · Sheet metal forming

✉ P. A. Prates
pedro.prates@dem.uc.pt

M. A. Dib
mario.dib@student.uc.pt

N. J. Oliveira
njoukov@student.dei.uc.pt

A. E. Marques
armando.marques@uc.pt

M. C. Oliveira
marta.oliveira@dem.uc.pt

J. V. Fernandes
valdemar.fernandes@dem.uc.pt

B. M. Ribeiro
bribeiro@dei.uc.pt

[1] Department of Informatics Engineering, Centre for Informatics and Systems of the University of Coimbra (CISUC), University of Coimbra, Coimbra, Portugal

[2] Department of Mechanical Engineering, Centre for Mechanical Engineering, Materials and Processes (CEMMPRE), University of Coimbra, Coimbra, Portugal

1 Introduction

Sheet metal forming is a manufacturing process that is commonly used for producing high-volume and low-cost components in the automotive, aircraft and home appliance industries. In this process, forces are applied to the metallic sheet to modify its geometry, enabling the production of complex shapes. The forces are applied by tools whose geometry dictates the shape of the component. The process design is complex because only the final shape of the component is known. Moreover, the process is highly nonlinear due to the large deformations imposed to the metal sheet, which presents plastic behaviour, but also as a result of the evolutionary boundary conditions imposed by the contact between the tools and the sheet. The conventional process design is based on empirical knowledge and an experimental "trial-and-error" approach. In this context, the virtual tryout of sheet metal forming components, based on the finite element method (FEM), has become an indispensable industrial tool to save design effort, money and time during the process set-up and production. The

rationale of FEM comes from the optimization of the process parameters, such as the tools geometry. FEM is a deterministic numerical tool since it enables the prediction of forming defects, such as localized necking, fracture and springback, for a predefined set of process and material parameters [4]. Nevertheless, it should be noticed that there are numerous variables involved in sheet metal forming process, which are related to the material properties, the tools geometry and process parameters. This makes the optimization of process conditions quite complex, particularly in the production of components which require several stages, and thus more than one set of tools. Therefore, the virtual tryout of sheet metal forming components with FEM is normally performed considering predefined material properties and values for some process parameters, such as the friction coefficient among others. In fact, the virtual tryout is still reliant on human expertise used to make key decisions at different stages of the design process. Still, even when resorting to FEM, unpredicted defects can occur in the experimental tryout or during production, which can be associated with the scatter observed in material properties, tools geometry and process parameters. The increasing competitiveness and relevance of sustainability issues in the industries lead to growing demands for high-quality components and reducing the costs generated by the production of defective components (scrap).

In this work, an approach to extract information from a sheet metal forming processes, exposed to sources of scatter in the material properties and process parameters, is proposed in order to enable the prediction of defects. The motivation is to reduce the costs and the time spent in the production of defective sheet metal components, i.e. contributing to improve the industry's efficiency. Machine Learning (ML) techniques are used, assuming that they can build models able to generalize well in unseen data. In this context, an empirical analysis of performance of ML techniques is conducted, considering single and ensemble classifiers. These are trained using data sets populated with numerical simulation results of two sheet metal forming processes: U-Channel and Square Cup. These processes were chosen for two main reasons: (i) they are benchmark tests commonly used to investigate the influence of the material as well as the process parameters on the occurrence of forming defects; (ii) they allow fast numerical simulation results, which is suitable for performing a large number of simulations. Since these processes present distinct features, different types of defects were considered for each one. Each type of defect is studied separately using a binary classification. Moreover, the data sets are generated for each forming process, for three steels with distinct mechanical properties.

The paper is organized as follows: Section 2 presents the details of the sheet metal forming processes and a review about ML applications in this context. Section 3 describes the proposed approach for evaluating the performance of both single and ensemble ML classifiers in predicting defects in sheet metal forming processes. The selected ML classifiers and ensemble methods are also discussed. Section 4 introduces the FEM models for the two forming processes under analysis. The procedure for generating and pre-processing the data sets as well as the evaluation metrics is also indicated. In Sect. 5, the results are presented and discussed. Firstly, the ML classifiers are analysed under a monolithic approach, considering also the influence of the size of sampling data. Afterwards, the analysis of the performance is conducted for the ensemble approach. Finally, the performance of both single and ensemble classifiers is compared. Section 6 presents the conclusions and future perspectives.

2 Background

2.1 Sheet metal forming

Sheet metal forming includes simple processes, such as bending, stretch forming and spinning, and more complex processes, like roll forming and deep drawing [16]. Each type of process has its specifications and parameters, including the tools geometry. Process design becomes even more complex when it is required to combine several processes and/or steps to produce the component. The main driver for the development of numerical tools, enabling the virtual tryout of sheet metal forming components, is the industry, in particular the automotive industry, due to the enormous amount of components involved in car production. The outer panels are usually the largest components, and their production involves the most complex operations, including deep drawing. As shown in Fig. 1, in deep drawing the metallic sheet is plastically deformed into the desired shape by the action of forming tools, which typically consist of a punch, a die and a blank holder. The blank (i.e. non-deformed metal sheet) is placed over the die, and it is forced to flow into the die cavity by the movement of the punch; the flow of the sheet is typically controlled with a blank holder, i.e. a tool that imposes a constant force on the flange region of the sheet. Thus, even for a simple forming process and assuming that the mechanical behaviour of the metallic sheet is known, there are many design variables to be considered, which are related to the blank and tools geometry and with the process control, such as the blank holder force. As previously mentioned, the FEM-based virtual tryout of sheet metal forming components enables a feasible process design to
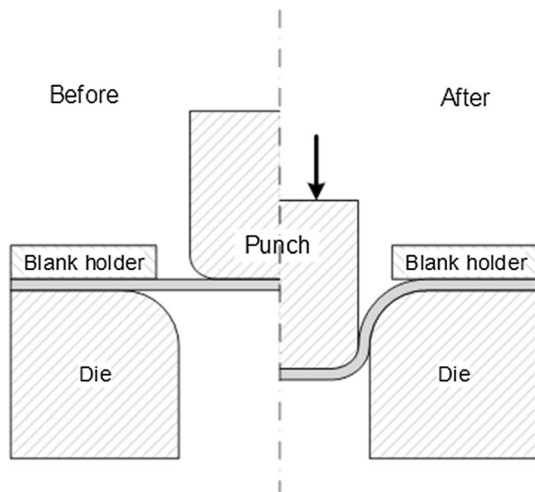
Fig. 1 Schematic representation of the sheet metal forming process

be achieved through the repetitive adjustment of process parameters based on the personal experience of the designer. However, to fully explore the finite element analysis, it has been combined with optimization algorithms in order to determine the process parameters automatically (e.g. [29, 37, 42]). This approach is more or less computationally intensive, depending on the number of design variables and the type of optimization algorithm selected. The number of trial experiments (i.e. numerical simulations) can be reduced resorting to a surrogate model (also called meta-model), used to guide the search for optimized parameter combinations. Different meta-modeling methods have been applied to the optimization of manufacturing process parameters (e.g. [18]), including artificial neural networks (ANNs) (e.g. [26, 34]). In the particular case of forming processes, researchers are also trying to explore the large amount of data generated (both experimental and numerical) while designing new products, to guide the process design from its early stage with the help of ANN meta-models to predict product feasibility (e.g. [40, 45]). In the context of early design stages, neural network classifiers have been applied in automating the sheet forming selection process, as an alternative to rule-based programmes [16]. Moreover, the robustness of the process design is questionable when neglecting the sources of scatter inherent to the process. The process design can be optimum for a specific combination of parameters but can easily lead to defective components due to slight variations introduced by scatter. In this context, a robust process window should be evaluated, in order to minimize the production of defective components (e.g. [46]). Therefore, much of the recent work has focused on statistical descriptions of variability within FEM, for assessing the sensitivity of defect predictions to the scatter of the parameters under analysis [19, 35, 47]. In FEM, the

material properties are commonly described using physics-based constitutive models. ML-based models have been pursued as an alternative to this type of models (e.g. [21, 23]), since the neural network does not require any prior assumption on the mathematical formulation between the input and output variables. The prime value added by ML is the ability to unveil the intrinsic response of a material in case of convoluted experimental data [24]. Nevertheless, some authors point out that physics-based constitutive models continue to provide useful insights to interpret the phenomena taking place, pursuing a different approach that uses machine learning to construct automatic corrections to existing models, based on data [14].

2.2 Machine learning applications to sheet metal forming

This subsection provides an overview of the literature on ML applications in sheet metal forming. Table 1 shows a comparative outline of ML applications in the prediction of forming defects [9, 13, 15, 20, 22, 25, 28, 30, 32, 39, 41]), which is the focus of the current work. Additional applications include: (i) material parameters' identification (e.g. [1–3, 6]); (ii) bend angles' prediction in laser forming processes (e.g. [7, 10]); (iii) die roll height prediction in fine blanking (e.g. [43, 48]); (iv) optimization of incremental sheet metal forming processes (e.g. [12, 17, 44]).

The summary presented in Table 1 highlights that ML applications have been focused on regression. In this regard, back-propagation-based artificial neural network (BP-ANN) is the primary option for the development of prediction models, some of them coupled with FEM analysis [9, 22, 28, 30, 32, 41]; ANN models trained with genetic algorithms (GA-ANN) were also developed [25, 39]. Most ML strategies in Table 1 are used to predict and account for springback in steel and aluminium parts obtained by sheet bending. This may be connected with the fact that springback (related to the elastic recovery of the material after tool release) is one of the main sources of geometrical and dimensional inaccuracy in sheet metal formed components, but also because of the simple geometries used. Nevertheless, models were also built to predict wrinkling and necking defects [9, 32]. In general, the features for training the ML models are material parameters (namely elastic and/or plastic properties) and the initial sheet thickness. This can be related to the fact that the standards for commercial metal sheets specify only a minimum allowable value or a relatively large range of values for the mechanical properties. Nevertheless, there are models that also consider process parameters. Globally, the literature review reveals that ML techniques to predict defects in sheet metal forming take into account different set-ups. Although promising results were reported,

Table 1 Summary of ML applications for the prediction of forming defects

Authors	Strategy	Forming process	Material	Features	Outputs
Inamdar et al. [20]	BP-ANN	Air V-bending	Steel and aluminium alloys (4)	Material parameters + Process parameters	Springback angle + Punch displacement
Guo and Tang [15]	BP-ANN	Air V-bending	Steel and aluminium alloys (5)	Sheet thickness + Material parameters + Process parameters	Bending springback angle
Miranda et al. [28]	ANN+FEM	Air V-bending	Steel alloys (2)	Sheet thickness + Process parameters	Punch displacement
Kazan et al. [22]	BP-ANN+FEM	Wipe bending	High-strength steel	Sheet thickness + Process parameters	Springback angle
Nasrollahi and Arezoo [30]	BP-ANN+FEM	Wipe bending on perforated metal sheets	Steel alloys (2)	Hole number and geometry + Process parameters + Type of material	Springback angle
Gisario et al. [13]	BP-ANN	Laser bending	Aluminium alloy	Starting deflection + Process parameters	Springback angle
Ruan et al. [39]	BP/GA-ANN	Multicurvature parts	Steel and aluminium alloys (4)	Sheet thickness + Process parameters	Springback angles
Liu et al. [25]	GA-ANN	U-bending	Not specified	Sheet thickness + Material parameters + Process parameters	Springback angle
Sharad and Nandedkar [41]	ANN+FEM	U-bending	Steel alloys (2)	Sheet thickness + Material parameters + Process parameters	Springback angles
Dib et al. [9]	MLP+FEM SVM+FEM DT+FEM RF+FEM NB+FEM	U-bending	Steel alloys (3)	Sheet thickness + Material parameters + Process parameters	Springback angle + Maximum thinning
Phatak et al. [32]	BP-ANN+FEM	Axisymmetric cup deep drawing	Aluminium alloy	Material parameters	Thickness + Friction coefficient

techniques to predict more than one type of defect for different types of materials and forming processes have rarely been considered. To the best of the authors' knowledge, there are currently no studies available in the literature regarding ML classification focused on defect prediction in sheet metal forming processes under variability, which is the main subject of the current work.

3 Proposed approach for building defect predictive models

This work focuses on the building of models able to predict the occurrence of defects for different types of materials and sheet metal forming processes under variability. Figure 2 presents the schematic diagram of the proposed approach, considering the branches for the monolithic and the ensemble classifiers. The first phase of both approaches consists in training the selected classifier. When resorting

to an ensemble model, either stacking or majority voting is used in the learning phase. Once the training phase is concluded, the predictive model is tested and the performance analysis is accomplished. To simplify Fig. 2, the predictive model is represented by only one box, although distinct types are built depending on the approach (monolithic or ensemble). To guarantee a proper comparison of performance, each model uses the same training and testing data, obtained from the same scaled sampling data. In addition, the same configuration with random weights was used.

3.1 Single learning classifier models

To accomplish the task for evaluating the best predictor of single sheet metal forming defects, seven ML classifiers were selected:

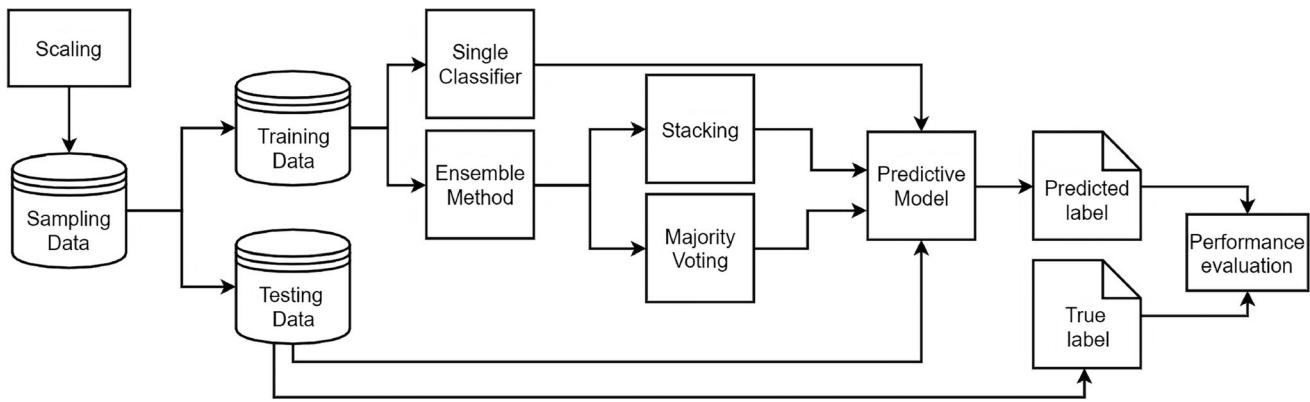– Multilayer perceptron (MLP)
– Decision tree (DT)

Fig. 2 Proposed approach for building sheet metal defect prediction models

- Random forest (RF)
- Naive Bayes (NB)
- Support vector machine (SVM)
- K-Nearest neighbours (KNN)
- Logistic regression (LR)

Seven ML models were created for each of the two types of defects considered in each of the two forming processes under analysis, for three different materials. The models were built using Python v3.6.2 and related libraries, such as SciPy Ecosystem and SciKit-learn, using default values for the parameters of each classifier [5, 33]. The following sections provide a theoretical background concerning each of the studied classifiers.

3.1.1 Multilayer perceptron (MLP)

The multilayer perceptron (MLP) is a class of feed-forward neural networks that consists of one input layer with n neurons $X_n = (x_1, x_2, ..., x_n)$, at least one hidden layer, where the number of hidden layers is arbitrary, and one output layer. Each layer has neurons that will connect with the neurons of the next layer, but they cannot be inter-connected. The MLP learning process is to adapt the connection weights in order to obtain a minimal difference between the network output and the desired output, resorting to learning algorithms such as back-propagation, which is based on gradient descent techniques.

The MLP output needs to compute the output of each unit in each layer, considering the set of hidden layers $H = (h_1, h_2, ..., h_n)$ and n_i neurons in each hidden layer h_i. The following equation is used to calculate the output of the first hidden layer:

$$h_i^l = \phi \left(\sum_j w_{ij}^{l-1} x_j + b_i^l \right), \tag{1}$$

where l is the layer position in the MLP architecture, ϕ is the activation function, that are nonlinear functions, w_{ij}^{l-1}

are the weights between the neuron i in the hidden layer $l - 1$ and the neuron j in the hidden layer $l + 1$. Finally, the network output is computed by:

$$y_i = \phi \left(\sum_j w_{ij}^l h_j^l + b_i^l \right), i = 1, ..., n \quad \text{and}$$

$$l = n \quad \text{and} \quad j = 1, ..., n, \tag{2}$$

where w_{ij}^l is the weight between the neuron i in the last hidden layer $l = n$, which is the output layer, and the neuron j in the output layer.

3.1.2 Decision tree (DT)

The decision tree (DT) is a nonparametric classifier that splits data continuously, based on simple decision rules. The choice of which feature to consider when splitting the data on each node is made in order to maximize information gain, which means minimizing:

$$F = \sum_{i=1}^{m} \frac{n_i}{N_i} H(D_i), \tag{3}$$

where n is the number of examples in the resulting node i with the desired label, N is the total number of examples in the resulting node i, D represents the data in said node, and H is an impurity function, such as entropy:

$$H(D) = -\sum_{i=1}^{m} p_i \log(p_i), \tag{4}$$

where p is the probability that an example in the data set corresponds to label i. This splitting process is repeated until each of the final nodes (leaves) only has samples with the same label. Alternatively, a stopping criterion can be defined in order to avoid overfitting.

3.1.3 Random forest (RF)

Random forest (RF) consists in a combination of several randomized decision trees and aggregating their predictions by averaging, characterizing the ensemble learning method, to solve classification and regression problems. In the binary supervised classification problem, the random response Y takes values in $\{0,1\}$ and a given input X has to guess the value of Y. A classification rule m_n is a measurable function of x and training sample T_n that attempts to estimate the label Y from x and T_n, where $T_n = (X_1, Y_1), ..., (X_n, Y_n)$ of independent random variables distributed the same as the independent prototype pair (X,Y), and $X = \{X_1, X_2, ..., X_n\}$.

A random forest is a predictor consisting of a collection of M randomized regression trees. For the kth tree in the family, the predicted value at the query point x is denoted by $m_n(x, \theta_k, T_n)$, where θ_k is a random vector generated with independent random variables of the kth tree, not related to the past random vectors $\theta_1, ..., \theta_{k-1}$ but with the same distribution.

In the classification situation, the random forest classifier is obtained via majority voting among the classification trees, that is:

$$m_{M,n}(x, \{\theta_1, ..., \theta_k\}, T_n)$$
$$= \begin{cases} 1 & \text{if} \quad \frac{1}{M}\sum_{k=1}^{M} m_n(x, \theta_k, T_n) > \frac{1}{2} \\ 0 & \text{Otherwise} \end{cases} \tag{5}$$

3.1.4 Naive Bayes (NB)

Naive Bayes is a classifier based on the application of the Bayes theorem, with the (naive) assumption that every pair of features is independent. Bayes' theorem states that:

$$P(y|x_1, ..., x_n) = \frac{P(y)P(x_1, ..., x_n|y)}{P(x_1, ..., x_n)} \tag{6}$$

After applying the naive assumption, this expression is simplified to:

$$P(y|x_1, ..., x_n) = \frac{P(y)\prod_{i=1}^{n} P(x_i|y)}{P(x_1, ..., x_n)} \tag{7}$$

For a given data set, the denominator will be the same for all entries, so a proportionality is considered:

$$P(y|x_1, ..., x_n) \propto P(y)\prod_{i=1}^{n} P(x_i|y) \tag{8}$$

The chosen label is the one that presents the maximum probability:

$$y = \text{argmax}_y P(y)\prod_{i=1}^{n} P(x_i|y) \tag{9}$$

3.1.5 Support vector machine (SVM)

Support vector machine (SVM) is supervised learning models, used to solve classification or regression problems. It is characterized as a discriminative classifier that finds the optimal separating hyperplane for test data points. The method consists in the binary classification of the training examples with features x and labels y, where $y \in \{-1, 1\}$, and uses the following function for classification:

$$h_{w,b}(x) = g(w^T x + b) \tag{10}$$

The SVM classifier directly predicts 1 or -1, instead of first estimating the probability of y being 1, where

$$g(z) = \begin{cases} 1 & \text{if} \quad z \geq 0 \\ -1 & \text{Otherwise} \end{cases} \tag{11}$$

The separating hyperplane is completely defined by (w, b). Given a training sample of (x_n, y_n), the functional margin can be defined as:

$$\hat{\gamma}_n = y_n(w^T x + b) \tag{12}$$

Given a training set that is linearly separable, the optimization problem described by the following equation should be solved:

$$\min_{\gamma, w, b} \quad \frac{1}{2}\|w\|$$
$$\text{s.t} \quad y_n = (w^T xn + b) \geq 1, i = 1, ..., m \tag{13}$$

The above is an optimization problem with a convex quadratic objective and only linear constraints, providing the optimal margin classifier.

3.1.6 K-nearest neighbours (KNN)

The k-nearest neighbours classifier does not create a model with the training data. Instead, each time it performs a prediction for a certain point, it starts by calculating the distance between each of the training data points and the test point. Then, the k training points that are nearest to the test point are selected, and these are used to make the prediction. The result of the prediction can be obtained by a simple majority vote from the selected training points. The KNN classifier is often known as a lazy learning since there is no training procedure but rather an assignment of the labels to the training instances in the first phase. In the second phase, the computation of the distance is performed as explained above.

3.1.7 Logistic regression (LR)

Logistic regression (LR) studies the association between a categorical dependent variable y and a set of independent (explanatory) variables x, where y consists of a binary code (0,1 or true, false) and x is numerical. With the requirements satisfied, this method fits a logistic curve, i.e. sigmoid curve, to the relationship between x and y. The sigmoid curve starts with slow, linear growth, followed by exponential growth, which then slows again to a stable rate. The simple logistic function is defined as follows:

$$y = \frac{e^x}{1 + e^x} \rightarrow y = \frac{1}{1 + e^{-x}} \qquad (14)$$

With the aim to provide more flexibility to the function, the logistic regression formula can be extended to a form where α and β are, respectively, the intercept of y and the regression coefficient:

$$y = \frac{e^{(\alpha + \beta x)}}{1 + e^{(\alpha + \beta x)}} \rightarrow y = \frac{1}{1 + e^{-(\alpha + \beta x)}} \qquad (15)$$

3.2 Ensemble models

Ensemble methods combine single classifiers, called base learners in this context, and are able to be more stable and predict better than single classifiers. The rationale is to reduce the bias and variance of the model to improve predictions. The goal is to build a model less noisy, more stable and less prone to overfitting. Since various base learners are used, each one can lead to a different prediction, where diversity among the base learners is a key aspect to ensemble performance. In this work, the following ensemble methods were used:

Majority Voting in the initial phase each base learner is trained. Afterwards, each base learner is fed with the testing data in order to have a prediction. The final predicted label is the one that has more than half of the votes;

Stacking similarly, in the initial phase, each base learner is trained. Afterwards, their outputs are used as features to train another ML classifier that is called meta-learner, which will make the final prediction.

Taking into account that majority voting favours the use of odd numbers of base learners, both methods were tested using combinations of 3 and 5 base learners. All possible combinations of the single classifiers described in Sect. 3.1 were tested. These classifiers were also tested as meta-learner. All this leads to a total of 56 combinations, for majority voting, and 392 combinations, for stacking.

4 Experimental set-up

4.1 Simulated data sets

The sampling data were generated using numerical simulation results, obtained with DD3IMP in-house FEM code [27, 31]. The numerical models for the U-Channel and the Square Cup processes are shown in Fig. 3. In both cases the total punch displacement is 30 mm. The U-Channel
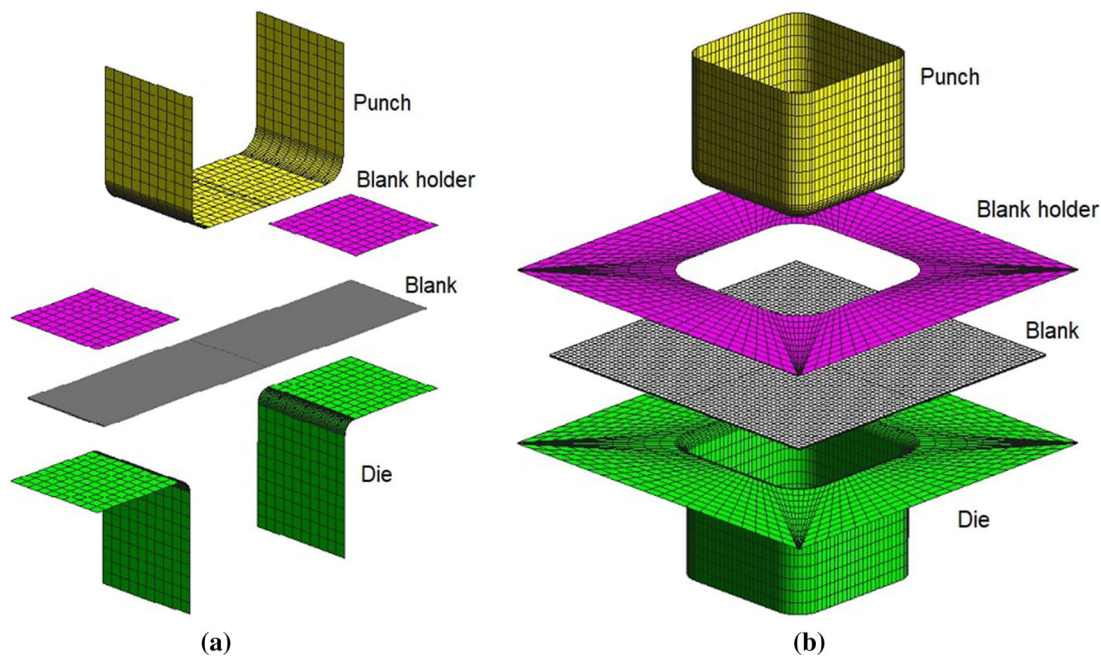

Fig. 3 Finite element models of: **a** U-Channel; **b** Square Cup

Table 2 Mean value and standard deviation of the material input features [35, 36]

Materials		C (MPa)	n	Y_0 (MPa)	E (GPa)	r_0	r_{45}	r_{90}	t_0 (mm)
DC06	μ	565.32	0.259	157.12	206	1.790	1.510	2.270	0.780
	σ	26.85	0.018	7.16	3.85	0.051	0.037	0.121	0.013
HSLA340	μ	673.00	0.131	365.30	210	0.820	1.070	1.040	0.780
	σ	32.30	0.011	10.67	7.35	0.033	0.039	0.061	0.005
DP600	μ	1093.00	0.187	330.30	210	1.010	0.760	0.980	0.780
	σ	52.46	0.020	9.64	7.35	0.040	0.030	0.060	0.010

For each material and feature, the first row is the mean value and the second row is the standard deviation. The Poisson ratio feature ($\mu = 0.3$, $\sigma = 0.015$) is identical for all materials

Table 3 Reference values for the non-defective U-Channel and Square Cup

Material	Springback (mm)		Maximum thinning (%)	
	BHF = 4.9 kN	BHF = 19.6 kN	BHF = 4.9 kN	BHF = 19.6 kN
U-Channel				
DC06	5.67	2.62	2.85	9.58
HSLA340	8.75	5.11	2.70	7.70
DP600	11.19	8.55	2.08	5.86
	Maximum EPS		Maximum thinning (%)	
	BHF = 2.45 kN	BHF = 9.8 kN	BHF = 2.45 kN	BHF = 9.8 kN
Square Cup				
DC06	0.92	0.87	14.01	20.00
HSLA340	0.92	0.86	24.22	47.43
DP600	1.14	0.85	28.35	39.18

corresponds to a bending process, and thus, the sheet is prone to significant springback. This type of defect is negligible in the Square Cup. In both cases, the occurrence of excessive thinning is an indicator of necking, which can also be controlled using the maximum equivalent plastic strain (EPS). Thus, two types of defects were considered: (i) springback and maximum thinning, for the U-Channel, and (ii) maximum EPS and maximum thinning, for the Square Cup. The tool geometry and the initial in-plane shape of the sheet are assumed fixed. Each process was simulated considering three steels commonly used in the automotive industry that cover a wide range of mechanical properties and applications: DC06 (mild steel), HSLA340 (high-strength low-alloy steel) and DP600 (dual-phase steel). The constitutive model considers: (i) elastic behaviour, Young's modulus, E, and Poisson ratio, v; (ii) plastic behaviour, yield stress, Y_0, strength and hardening coefficients, C and n, and anisotropy coefficients r_0, r_{45} and r_{90}. The initial sheet thickness t_0 is also considered. The variability in the input features related to the material parameters is typified by a normal distribution, with mean (μ) and standard deviation (σ) values shown in Table 2.

Two input features related to process parameters were also considered: the friction coefficient and the blank holder force (BHF). The mean value of the friction coefficient is 0.144 for all materials, with $\sigma/\mu = 20\%$ [19]. For

the BHF, two mean values were considered, which correspond to a lower and an upper level of the process window. For the U-Channel, the mean values used were 4.9 and 19.6 kN, while for the Square Cup, they were 2.45 and 9.8 kN. For each BHF value, the variability is $\sigma/\mu = 5\%$. Thus, the variability of a total of 11 features was considered in analysis of both forming processes, for the three materials. In this context, random numerical simulations were performed within the range of variation of the input features (see Table 2). The numerical simulations using the mean values of the input features presented in Table 2 lead to a non-defective solution, which is considered as a reference solution. A defect occurs when the output value obtained from the random simulations is greater than that of the reference solution, whose values are presented in Table 3.

4.2 Data set pre-processing

The data sets were split in training (70%) and testing data (30%). Data scaling was performed in all the data sets. A maximum of 2000 experiments (i.e. numerical simulation results) was considered for each material and forming process. The data were randomly shuffled in order to repeat the process 30 times (runs).
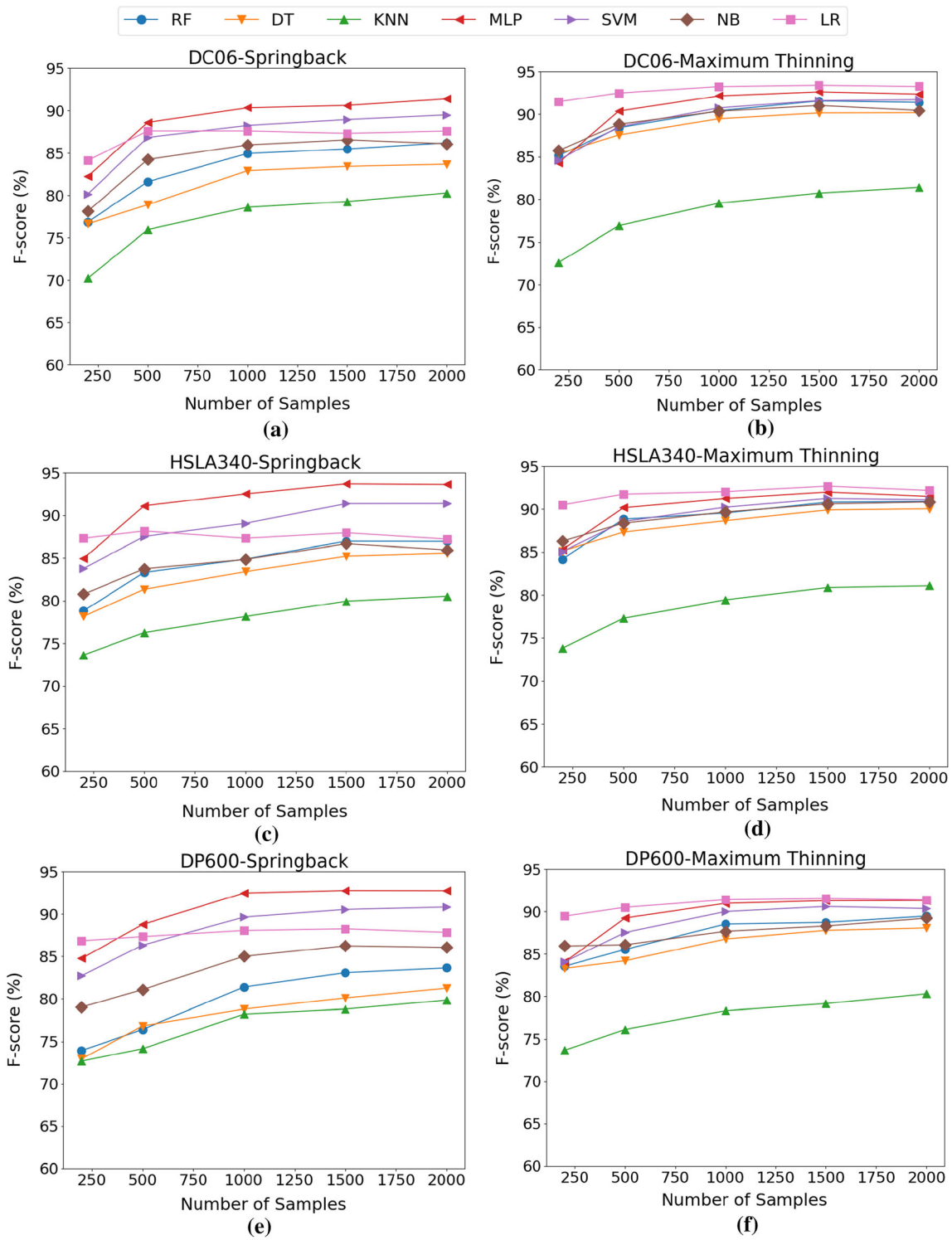
Fig. 4 F-score per number of samples for the U-Channel predictive models

4.3 Performance measures

The F-score measure is used to evaluate the performance of both single and ensemble classifiers. This performance metric combines both precision and recall metrics and provides a break-even between them. The F-score is calculated as the harmonic mean between precision and recall, as follows:
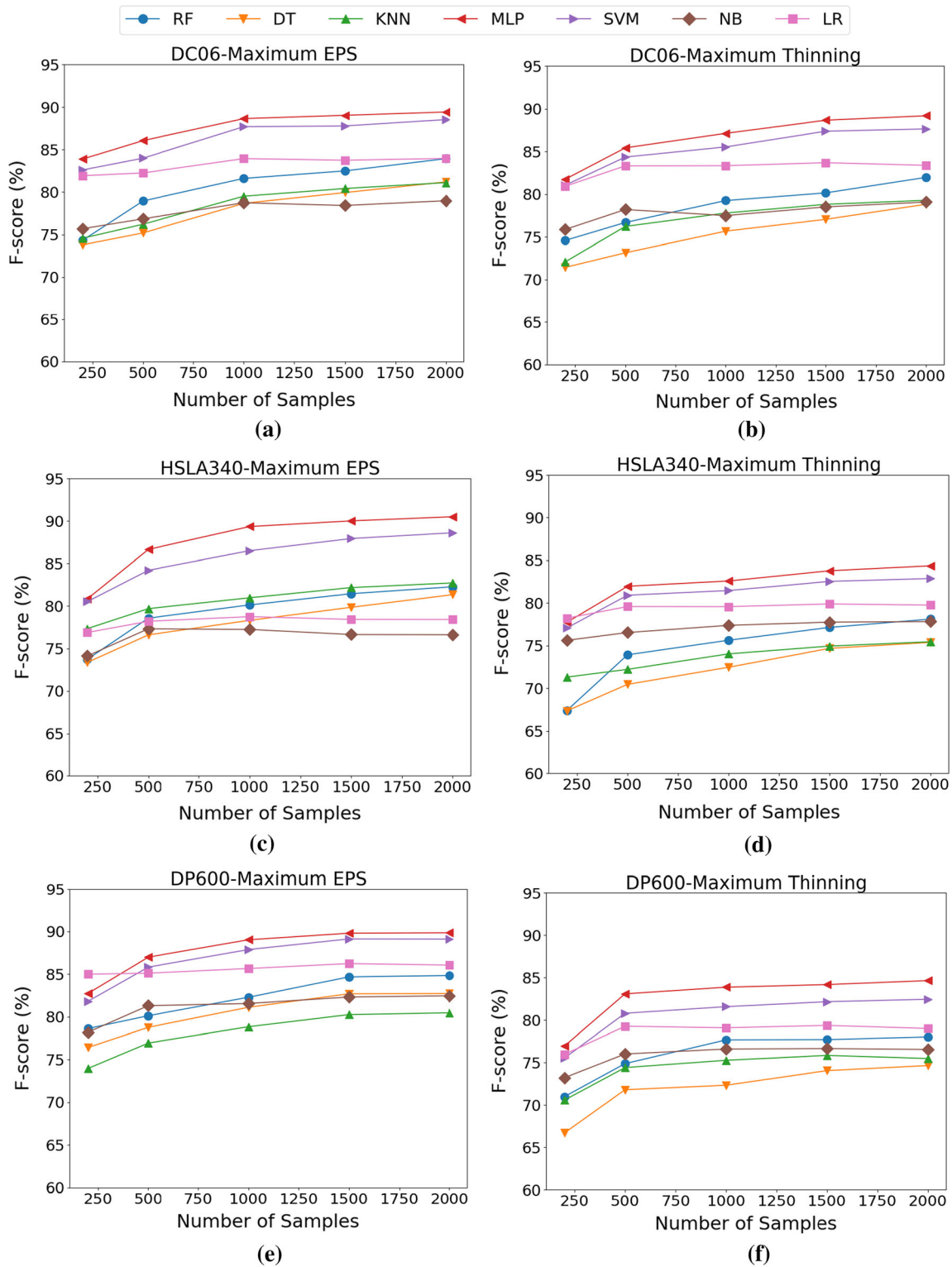
Fig. 5 F-score per number of samples for the Square Cup predictive models

$$F\text{-score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}, \qquad (16)$$

where precision takes into account the proportion of

correctly classified instances (true positives (TPs)), among all the positive instances classified (true positives (TP) and false positives (FP)),
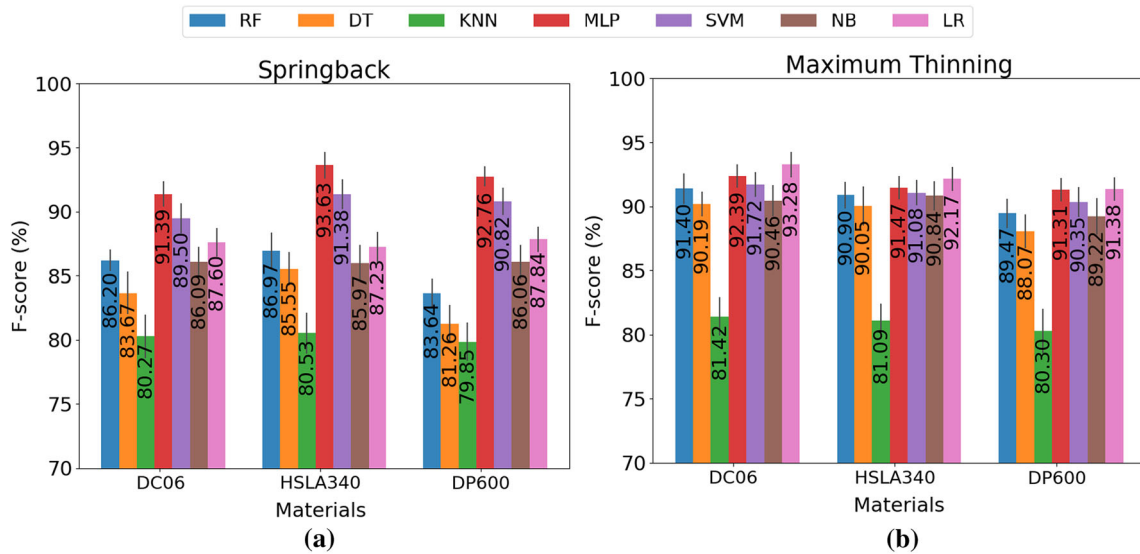
Fig. 6 F-score of the classifiers for the distinct materials and defects regarding the U-Channel predictive models

$$\text{Precision} = \frac{TP}{TP + FP}, \tag{17}$$

and recall evaluates the percentage of correctly identified instances of a class (TP) among all the instances of a given class (true positives (TP) and false negatives (FN)):

$$\text{Recall} = \frac{TP}{TP + FN} \tag{18}$$

5 Results and discussion

5.1 Single classifiers

Figures 4 and 5 show the evolution of the F-score values with the sampling data size, respectively, for the U-Channel and Square Cup forming predictive models, using 200, 500, 1000, 1500 and 2000 samples. In general, for both the U-Channel and Square Cup, the values of F-score increase with the increase in the sampling data size; exceptions include the cases "HSLA340-springback" and "DP600-springback" with the LR classifier (see Fig. 4c, e), where the F-score is nearly constant. Accordingly, the highest values of F-score are generally obtained for 2000 samples, with few exceptions. Adding more training data would reduce variance but increase bias. Therefore, the performance analysis will focus on the results with 2000 samples, which is considered the critical sampling size for this problem [38]. Figure 6 shows the values of F-score, for the critical sampling size, obtained by the U-Channel predictive models in the cases of springback (Fig. 6b) and maximum thinning (Fig. 6c). The mean and standard deviation values of the F-score were obtained from 30 runs of each

single classifier. The mean values of F-score range from 79.85% (springback prediction with DP600, using KNN— see Fig. 6b) to 93.63% (springback prediction with HSLA340, using MLP—see Fig. 6b), with relatively low standard deviation values. MLP is the highest performing classifier for predicting springback, with mean F-score values of 91.39% (DC06), 93.63% (HSLA340) and 92.76% (DP600), and LR is the highest performing classifier for predicting maximum thinning, with mean F-score values of 93.28% (DC06), 92.17% (HSLA340) and 91.38% (DP600). On the other hand, KNN is the lowest performing classifier, with mean values of F-score close to 80% for all materials and both defects (see Fig. 6b, c). The dissimilarity between the performances of the classifiers is more noticeable for the springback (Fig. 6b) than for the maximum thinning (Fig. 6c), in which all the classifiers except KNN are competitive.

Figure 7 shows the values of F-score, for the critical sampling size, obtained by the Square Cup predictive models of maximum EPS (Fig. 7b) and maximum thinning (Fig. 7c). The mean values of F-score range from 74.65% (maximum thinning prediction with DP600, using DT—see Fig. 7c) to 90.50% (maximum EPS prediction with HSLA340, using MLP—see Fig. 7b), with relatively low standard deviation values. The MLP is the highest performing classifier for predicting both the maximum EPS and maximum thinning in all materials, with mean values of F-score ranging from 84.37% (maximum thinning, HSLA340) to 90.50% (maximum EPS, HSLA340); also, the SVM classifiers show a relatively good performance. The NB, KNN and DT classifiers are the lowest performing classifiers.
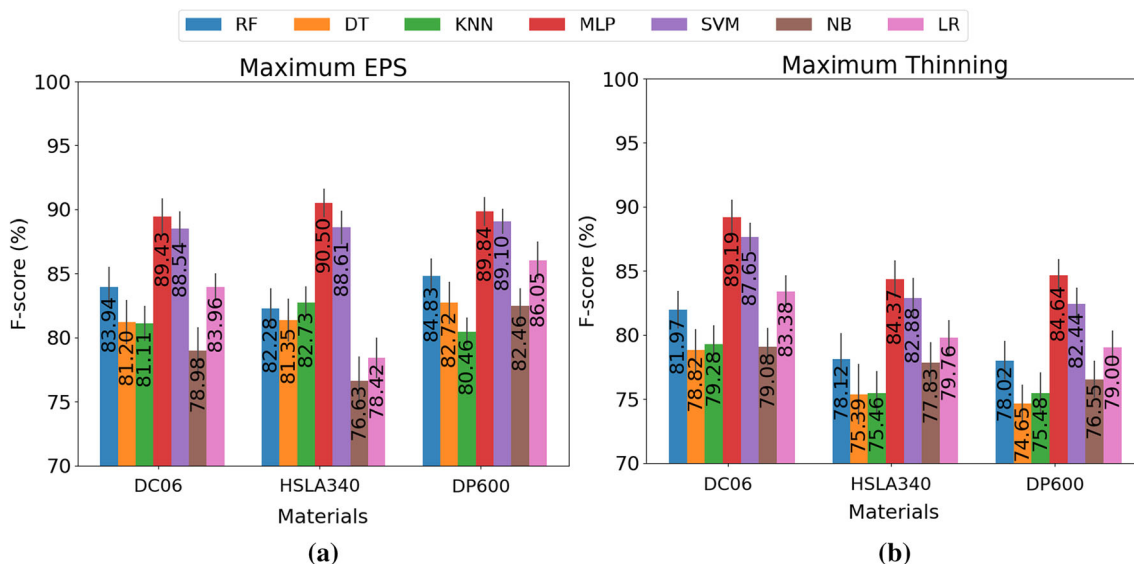
Fig. 7 F-score of the classifiers for the distinct materials and defects regarding the Square Cup predictive models

The results show that, for a given type of defect, most classifiers show similar performances among the materials. For a given material, the difference in the performance of the classifiers between the two types of defects is more noticeable in the U-Channel than in the Square Cup. It is further noticed that MLP and KNN are, respectively, the highest and the lowest performing classifiers. This indicates that learning is an important step for finding the nonlinear decision boundaries. In fact, KNN is a "lazy learner" and it is harder to discriminate between the sought classes; thus, as a predictor, it becomes less useful. Finally, a Friedman test was conducted on the respective F-score values of each classifier, to check whether the performances of the single classifiers are significantly different; this nonparametric statistical test allows for performance comparison when dealing with several classifiers over multiple data sets [8, 11]. In this test, the null hypothesis states that all single classifiers performed equally. The rejection of this hypothesis means that differences between the performances of single classifiers are statistically significant. The obtained Friedman statistic, equal to 59.82 (i.e. corresponding to a p value equal to 4.89×10^{-11}), is greater than its critical values at significance levels of 5% (12.59) and 1% (16.81), which lead us to reject the null hypothesis.

5.2 Ensemble classifiers

Tables 4 and 5 present the best combinations of the classifiers obtained for majority voting and stacking ensembles, respectively. The mean and standard deviation values of the F-score were obtained from 30 runs of each ensemble. When comparing both tables, the use of stacking ensembles generally leads to an increase in the performance relatively

to majority voting; the increase in performance corresponds to more than 1.5% in the case of the Square Cup (see cases "HSLA340—maximum EPS" and "DC06—maximum thinning" in Table 5). The opposite occurs only for the maximum thinning prediction in the U-Channel (see cases "U-Channel—Maximum Thinning" in Tables 4 and 5), where a maximum performance reduction in 0.4% is obtained for DP600. In the case of stacking ensembles, the area under the curve (AUC) metric was determined (see Table 5), which depicts the trade-off between the true-positive rate and the false-positive rate. The relatively high values of AUC (generally above 90% average

Table 4 F-score for the majority voting ensembles

Material	Algorithms	F-score (%)
U-Channel–springback		
DC06	DT, MLP, SVM	90.82 ± 1.07
HSLA340	DT, MLP, SVM	92.90 ± 0.93
DP600	DT, MLP, SVM	92.17 ± 0.93
U-Channel–maximum thinning		
DC06	DT, MLP, LR	93.42 ± 0.75
HSLA340	RF, MLP, LR	92.24 ± 0.90
DP600	DT, MLP, LR	91.98 ± 1.17
Square Cup–maximum EPS		
DC06	RF, DT, KNN, MLP, SVM	88.45 ± 1.26
HSLA340	KNN, MLP, SVM	88.87 ± 1.19
DP600	RF, DT, KNN, MLP, SVM	89.58 ± 1.47
Square Cup–maximum thinning		
DC06	RF, DT, MLP, SVM, LR	87.62 ± 1.39
HSLA340	DT, MLP, SVM	83.58 ± 1.35
DP600	DT, MLP, SVM	83.80 ± 1.54

Table 5 F-score and AUC for the stacking ensembles

Material	Algorithms		F-score (%)	AUC (%)
	Base learners	Meta-classifier		
U-Channel–springback				
DC06	KNN, MLP, SVM	SVM	91.42 ± 1.04	93.53 ± 1.12
HSLA340	KNN, MLP, SVM	SVM	93.74 ± 0.88	95.65 ± 0.79
DP600	KNN, MLP, SVM	SVM	92.89 ± 0.75	95.33 ± 0.71
U-Channel–maximum thinning				
DC06	KNN, NB, LR	SVM	93.25 ± 0.98	95.29 ± 1.18
HSLA340	RF, DT, MLP, SVM, LR	MLP	92.09 ± 1.16	96.38 ± 0.68
DP600	RF, DT, MLP, SVM, LR	MLP	91.58 ± 1.13	96.25 ± 0.81
Square Cup–maximum EPS				
DC06	MLP, SVM, LR	RF	89.74 ± 1.37	92.22 ± 1.48
HSLA340	KNN, MLP, SVM	SVM	90.49 ± 0.98	91.20 ± 1.84
DP600	MLP, SVM, LR	LR	90.01 ± 1.08	92.96 ± 0.96
Square Cup–maximum thinning				
DC06	KNN, MLP, SVM	RF	89.33 ± 1.45	91.87 ± 1.26
HSLA340	KNN, MLP, NB	RF	84.47 ± 1.39	87.32 ± 1.49
DP600	KNN, MLP, SVM	LR	84.62 ± 1.34	87.22 ± 1.09

performance) and their low variance (ca. 1%) confirm the high robustness of the stacking ensembles.

The smallest standard deviation values were obtained in both springback and maximum thinning U-Channel ensembles, respectively, stacking for DP600 (see Table 5) and majority voting for DC06 (see Table 4). The low values of standard deviations on all the 30 runs in all the experiments reveal a great deal of stability in the procedure as well as relatively good significance of the results.

The performance comparison between single and ensemble classifiers shows that the latter provide generally better defects predictors. In particular, in stacking ensembles it is expected that the meta-learner is prone to less errors by reducing error variance and thus generalizing well in the test set. Ensemble methods combine several machine learning techniques into one predictive model in order to decrease variance and bias, or improve predictions. The fact that the single classifiers already provide relatively high performances (ca. 85%, on average) and thus are stable learners; the best combinations of the classifiers obtained for the majority voting and stacking ensembles (ca. 90% average performance) do not show to outperform significantly the single classifiers. On the other hand, the overall increase in the ensembles' performance is coupled with a lower variance, which promotes robustness and stability of the procedure and allows a better bias–variance trade-off.

6 Conclusion

In this study, machine learning techniques were used for predicting defects of sheet metal forming processes. The same sampling data were applied to generate single learning and ensemble models. These data were obtained using numerical simulations of two forming processes, U-Channel and Square Cup, with three different materials. In general, the performance of single classifiers increases with the increase in the sampling data size, showing a stabilization that enables the definition of a critical sampling size. Considering the critical sampling size, the results show that, for a given type of defect, most single classifiers show similar performances among the materials. The best combinations of the classifiers obtained for the majority voting and stacking ensembles can provide better predictors than single classifiers (particularly when using stacking ensembles); however, the performance differences are small. Ensemble models allow a better trade-off between bias and variance, and it is expected they perform well in real data from sheet metal forming industry. In fact, the relatively high F-score values coupled with their low variance motivate the application of the proposed approach in industrial environment, in order to assess its feasibility as a decision support tool for predicting defects in sheet metal forming. Further studies will focus on the development of ML regression models for predicting sheet metal forming defects, and the subsequent performance comparison with response surface methodology and kriging regression models.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Abbassi F, Belhadj T, Mistou S, Zghal A (2013) Parameter identification of a mechanical ductile damage using artificial neural networks in sheet metal forming. Mater Des 45:605–615
2. Aguir H, Chamekh A, Belhadjsalah H, Dogui A, Hambli R (2008) Identification of constitutive parameters using hybrid ANN multi-objective optimization procedure. Int J Mater Form 1:1–4
3. Aguir H, BelHadjSalah H, Hambli R (2011) Parameter identification of an elasto-plastic behaviour using artificial neural networks-genetic algorithm method. Mater Des 32(1):48–53
4. Banabic D (2010) Sheet metal forming processes: constitutive modelling and numerical simulation. Springer, Berlin
5. Brownlee J (2016) Machine learning mastery with Python: understand your data, create accurate models and work projects end-to-end, 1st edn. Jason Brownlee, Melbourne
6. Chamekh A, Bel Hadj Salah H, Hambli R (2008) Inverse technique identification of material parameters using finite element and neural network computation. Int J Adv Manuf Technol 44(1):173
7. Cheng P, Lin S (2000) Using neural networks to predict bending angle of sheet metal formed by laser. Int J Mach Tools Manuf 40:1185–1197
8. Demšar J (2006) Statistical comparisons of classifiers over multiple data sets. J Mach Learn Res 7:1–30
9. Dib M, Ribeiro B, Prates P (2018) Model prediction of defects in sheet metal forming processes. In: Pimenidis E, Jayne C (eds) Eng Appl Neural Netw. Springer, Cham, pp 169–180
10. Fetene BN, Shufen R, Dixit US (2016) Fem-based neural network modeling of laser-assisted bending. Neural Comput Appl 29:69–82
11. Friedman M (1937) The use of ranks to avoid the assumption of normality implicit in the analysis of variance. J Am Stat Assoc 32(200):675–701. https://doi.org/10.1080/01621459.1937.10503522
12. Fu Z, Mo J, Chen L, Chen W (2010) Using genetic algorithm-back propagation neural network prediction and finite-element model simulation to optimize the process of multiple-step incremental air-bending forming of sheet metal. Mater Des 31(1):267–277
13. Gisario A, Barletta M, Conti C, Guarino S (2011) Springback control in sheet metal bending by laser-assisted bending: experimental analysis, empirical and neural network modelling. Opt Lasers Eng 12:1372–1383
14. González D, Chinesta F, Cueto E (2019) Learning corrections for hyperelastic models from data. Front Mater 6:14. https://doi.org/10.3389/fmats.2019.00014
15. Guo Z, Tang W (2017) Bending angle prediction model based on BPNN-spline in air bending springback process. Math Probl Eng 2017:11
16. Hamouche E, Loukaides EG (2018) Classification and selection of sheet forming processes with machine learning. Int J Comput Integr Manuf 31:921–932
17. Hartmann C, Opritescu D, Volk W (2019) An artificial neural network approach for tool path generation in incremental sheet metal free-forming. J Intell Manuf 30(2):757–770. https://doi.org/10.1007/s10845-016-1279-x
18. Havinga G (2016) Optimization and control of metal forming processes. PhD thesis, University of Twente. https://doi.org/10.3990/1.9789036541480
19. Huang C, Radi B, Hami A (2016) Uncertainty analysis of deep drawing using surrogate model based probabilistic method. Int J Adv Manuf Technol 86:9–12
20. Inamdar M, Date P, Desai U (2000) Studies on the prediction of springback in air vee bending of metallic sheets using an artificial neural network. J Mater Process Technol 108:45–54
21. Jenab A, Sarraf IS, Green DE, Rahmaan T, Worswick MJ (2016) The use of genetic algorithm and neural network to predict rate-dependent tensile flow behaviour of AA5182-O sheets. Mater Des 94:262–273. https://doi.org/10.1016/j.matdes.2016.01.038
22. Kazan R, Firat M, Tiryaki AE (2007) Prediction of springback in wipe-bending process of sheet metal using neural network. Mater Des 30:418–423
23. Kessler BS, El-Gizawy AS, Smith DE (2006) Incorporating neural network material models within finite element analysis for rheological behavior prediction. J Press Vessel Technol 129:58–65. https://doi.org/10.1115/1.2389004
24. Li X, Roth CC, Mohr D (2019) Machine-learning based temperature- and rate-dependent plasticity model: application to analysis of fracture experiments on DP steel. Int J Plast 118:320–344. https://doi.org/10.1016/j.ijplas.2019.02.012
25. Liu W, Liu Q, Ruan F, Liang Z, Qiu H (2007) Springback prediction for sheet metal forming based on GA-ANN technology. J Mater Process Technol 187–188:227–231
26. Mekras N (2017) Using artificial neural networks to model aluminium based sheet forming processes and tools details. J Phys Conf Ser 896:012090. https://doi.org/10.1088/1742-6596/896/1/012090
27. Menezes LF, Teodosiu C (2000) Three-dimensional numerical simulation of the deep-drawing process using solid finite elements. J Mater Process Technol 97:100–106
28. Miranda SS, Barbosa MR, Santos AD, Pacheco JB, Amaral RL (2018) Forming and springback prediction in press brake air bending combining finite element analysis and neural networks. J Strain Anal Eng Des 53(8):584–601
29. Naceur H, Ben-Elechi S, Batoz J, Knopf-Lenoir C (2008) Response surface methodology for the rapid design of aluminum sheet metal forming parameters. Mater Des 29(4):781–790
30. Nasrollahi V, Arezoo B (2012) Prediction of springback in sheet metal components with holes on the bending area, using experiments, finite element and neural networks. Mater Des 36:331–336
31. Oliveira MC, Alves JL, Menezes LF (2008) Algorithms and strategies for treatment of large deformation frictional contact in the numerical simulation of deep drawing process. Arch Comput Methods Eng 15(2):113–162

32. Pathak K, Anand VK, Agnihotri G (2008) Prediction of geometrical instabilities in deep drawing using artificial neural network. J Eng Appl Sci 3:344–349

33. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J, Passos A, Cournapeau D, Brucher M, Perrot M, Duchesnay E (2011) Scikit-learn: machine learning in Python. J Mach Learn Res 12:2825–2830

34. Pfrommer J, Zimmerling C, Liu J, Kärger L, Henning F, Beyerer J (2018) Optimisation of manufacturing process parameters using deep neural networks as surrogate models. Procedia CIRP 72:426–431. https://doi.org/10.1016/j.procir.2018.03.046

35. Prates PA, Adaixo AS, Oliveira MC, Fernandes JV (2018) Numerical study on the effect of mechanical properties variability in sheet metal forming processes. Int J Adv Manuf Technol 96:561–580

36. Prates PA, Marques AE, Oliveira MC, Fernandes JV (2019) Comparing metamodeling techniques for variability analysis in sheet metal forming processes. AIP Conf Proc 2113(1):120003. https://doi.org/10.1063/1.5112658

37. Qiuchong Z, Yuqi L, Zhibing Z (2016) A new optimization method for sheet metal forming processes based on an iterative learning control model. Int J Adv Manuf Technol 85(5):1063–1075

38. Ribeiro B, Silva J, Sung A (2018) Critical feature selection and critical sampling for data mining. In: G. Ganapathi (eds.), 3rd international conference on computational intelligence, cyber security & computational models. Springer Nature Singapore, ICC3, CCIS 844, pp 13–24

39. Ruan F, Feng Y, Liu W (2008) Springback prediction for complex sheet metal forming parts based on genetic neural network. In: 2008 second international symposium on intelligent information technology application

40. Sauer C, Schleich B, Wartzack S (2018) Deep learning in sheet-bulk metal forming part design. In: DS 92: Proceedings of the DESIGN 2018 15th international design conference, pp 2999–3010

41. Sharada G, Dr Nandedkar VM (2014) Springback in sheet metal u bending-FEA and neural network approach. Procedia Mater Sci 6:835–839

42. Shi X, Chen J, Peng Y, Ruan X (2004) A new approach of die shape optimization for sheet metal forming processes. J Mater Process Technol 152:35–42

43. Stanke J, Feuerhack A, Trauth D, Mattfeld P, Klocke F (2018) A predictive model for die roll height in fine blanking using machine learning methods. Procedia Manufacturing 15:570 – 577. In: Proceedings of the 17th international conference on metal forming METAL FORMING 2018 September 16–19, 2018, Loisir Hotel Toyohashi, Toyohashi, Japan

44. Stoerkle DD, Seim P, Thyssen L, Kuhlenkoetter B (2016) Machine learning in incremental sheet forming. In: Proceedings of ISR 2016: 47st international symposium on robotics, pp 1–7

45. Wessing S, Rudolph G, Turck S, Klimmek C, Schäfer SC, Schneider M, Lehmann U (2014) Replacing FEA for sheet metal forming by surrogate modeling. Cogent Eng 1(1):950853. https://doi.org/10.1080/23311916.2014.950853

46. Wiebenga J, Atzema E, van den Boogaard A (2015) Stretching the limits of forming processes by robust optimization: a numerical and experimental demonstrator. J Mater Process Technol 217:345–355

47. Wiebenga JH, Atzema EH, An YG, Vegter H, Boogaard AH (2014) Effect of material scatter on the plastic behavior and stretchability in sheet metal forming. J Mater Process Technol 214(2):238–252

48. Zhuang X, Zhang W, Wu Y, Zhao Z (2018) Comprehensive prediction method for die-roll height of fine-blanking components. Int J Adv Manuf Technol 98(9):2819–2829