APEM journal
Advances in Production Engineering & Management
Volume 19 | Number 1 | March 2024 | pp 78–92
https://doi.org/10.14743/apem2024.1.494
ISSN 1854-6250
Journal home: apem-journal.org
Original scientific paper

A comparative study of machine learning regression models for production systems condition monitoring

Jankovič, D.[a], **Šimic, M.**[a,*], **Herakovič, N.**[a]

[a]University of Ljubljana, Faculty of Mechanical Engineering, Department of Manufacturing Technologies and Systems, Ljubljana, Slovenia

A B S T R A C T

This research investigates the benefits of different Machine Learning (ML) approaches in production systems, with respect to the given use case of considering the forming process and different friction conditions on hydraulic press response in between the phases of the sheet metal bending cycle, i.e. bending, levelling and movement. A framework for enhancing production systems with ML facilitates the transition to smarter processes and enables fast, accurate predictions integrated into decision-making and adaptive control. Comparative ML analysis provides insights into predictive regression models for hydraulic press condition recognition, enhancing process improvement. Our results are supported by performance evaluation metrics of predictive accuracy RMSE, MAE, MSE and R^2 for Linear Regression (LR), Decision Trees (DT), Support Vector Machine (SVM), Gaussian Process Regression (GPR) and Neural Network (NN) models. Given the remarkable predictive accuracy of the regression models with R^2 values between 0.9483 and 0.9995, it is noteworthy that less complex models exhibit significantly shorter training times, up to 437 times shorter than more complex models. In addition, simpler models have up to 36 times better prediction rates, compared to more complex models. The fundamentals illustrate the trade-offs between model complexity, accuracy and computational training and prediction rate.

A R T I C L E I N F O

Keywords:
Hydraulic press;
Metal forming;
Machine Learning (ML);
Linear Regression (LR);
Decision Trees (DT);
Support Vector Machine (SVM);
Gaussian Process Regression (GPR);
Artificial Neural Networks (ANN)

**Corresponding author:*
marko.simic@fs.uni-lj.si
(Šimic, M.)

Article history:
Received 1 February 2024
Revised 19 April 2024
Accepted 23 April 2024

1. Introduction

Challenges in manufacturing processes often require straightforward solutions or are so complex that they present technicians with seemingly unsolvable problems. Simulation approaches are often used to overcome these challenges in order to gain a more detailed understanding of the system under consideration and the manufacturing processes running on it [1]. However, the underlying problems are sometimes indirect when comparing the real environment with the virtual environment established in a simulation model [2]. With the advent of Industry 4.0, technological solutions have emerged in digitalization and automation processes, often using artificial intelligence (AI) approaches [3]. The synergy of edge computing (EC) and 5G networks leverages high bandwidths and low latency to process data closer to the source, enabling faster response times and more efficient, stable and secure data transfer and data management, which has a positive impact on the effectiveness and quality of production processes as well as sustainability and circular economy improvement [4]. The ML paradigm enables a more in-depth system analysis that captures properties that cannot be evaluated by visual representation and graphical interpretation alone. Classification and regression models based on data-driven approaches have already been successfully implemented in many technical fields, offering high precision in predicting system behaviour and often overcoming 99 % prediction efficiency in real systems [5, 6].

Hydraulic presses play a crucial role in various forming processes such as bending, stamping, forging and drawing, where the aim is to convert hydraulic energy into deformation energy of the workpiece. The dynamic properties of hydraulic presses and bending processes, including material properties, fluid properties and friction behaviour, are interrelated and influence each other. These complex relationships require in-depth investigations, such as those presented in this study, to gain a comprehensive understanding of how the bending process affects the response of the hydraulic press. The quality of the forming process has a significant impact on the overall quality of the product, especially when disturbances, such as dynamic frictional properties, are present [7, 8]. Identifying altered conditions is a fundamental step in comprehending system behaviour, as highlighted in prior research on issues like hydraulic valve wear [9], pump malfunction [10] and leakage [11] has shown. To summarise, artificial intelligence offers data-driven modelling solutions tailored to different engineering challenges and based on different principles. Briefly speaking, ML provides solutions for the identification of faults and malfunctions of hydraulic components, which are crucial for decision making to adapt and improve the operation of hydraulic presses.

A detailed analysis in the field of hydraulic systems shows that there is no simple method for selecting the most appropriate ML approach, regardless of the input parameters. Given the use of numerous classification and regression methods in studies, a significant dilemma arises when choosing between less complex methods such as LR, DT, SVR and more complex methods such as GPR and NN for the study of specific research problems [12–17]. Another possible shortcoming of the previously referenced research refers to the fact the authors believe that certain derivations of ML methods are best suited for their particular use case due to their complexity. This means that the optimization and prediction of performance depends on the method chosen for the particular application, which requires careful consideration. Su *et al.* [18]demonstrated that when comparing the ML methods LR, K Nearest Neighbour (KNN), DT, SVM and NN, the NN method outperformed the others with an accuracy of 99.8 % in predicting valve flow. Highlighting the best ML method, it is worth noting that other ML models achieved robust prediction accuracy, with the worst performing LR model still achieving 99.1 %. Moreover, the same conclusions have been confirmed by research groups investigating other use cases [19–22]. On the other hand, Guo *et al.* [23] emphasize that training and prediction times are a crucial parameter for the efficient integration of ML models into decision algorithms. It determines how quickly the entire decision-making system will react to changing conditions and therefore allows real-time actions such as control and parameters set-up.

By pointing out the basic assumptions and an analysis of the available ML algorithms, the objective of this paper is to demonstrate the advantages of the five basic regression modelling types LR, DT, SVM, GPR, NN in the investigation of a hydraulic press under different intensities of the forming process and friction dynamics. Chapter 2 introduces the background and the research problem, supplemented by Design of Experiments (DOE), data extraction and data preprocessing. Section 3 introduces different types of regression models and defines the range of hyperparameters to obtain the best prediction from five ML approaches. Finally, Section 4 presents the best fitted regression models along with the most efficient selection of hyperparameters.

2. Experiment and dataset analysis

2.1 A case study of hydraulic press

This study focuses on a hydraulic press subjected to the conditions of the forming process, including the bending process and the constraints on the movement of the hydraulic cylinder resulting from the friction between the press guides and the hydraulic cylinder, as shown in Fig. 1. A hydraulic press is a sophisticated and robust mechanical device designed for various industrial applications that utilises the principles of fluid mechanics to generate significant force for forming, shaping and compressing materials. The design of a hydraulic press combines

- hydraulic components, i.e. hydraulic valve, hydraulic cylinder, hydraulic power unit (hydraulic pump with electric motor and controller),
- mechanical components, i.e. guiding system, hydraulic press frame, mechanical system (bending process),
- control system, e.g. PLC Beckhoff Controller CX 9020, PID Moog Controller, Raspberry Pi.

The components listed above with their respective characteristics form a distributed network in which data collection and analysis takes place. They therefore form the basis for integration into the concept of edge computing and 5G communication technology and particular on 5G real-time data transfer. The integration of these components enables the hydraulic press to effectively convert hydraulic energy into mechanical power. As explained by Jankovic D. et. Al [24], these influences contribute to a delayed movement of the hydraulic cylinder, which is determined not only by the mentioned factors, but also by the intensity of these factors and the velocity of the hydraulic cylinder. To achieve this goal, 30 different scenarios were analysed in the experiments, focusing on the effects of the forming process. In addition, a further 46 scenarios were systematically carried out to evaluate the influence of different operating conditions on the reaction of the hydraulic press.

Forming processes require the mechanical energy of a hydraulic press to change the shape of the specimen. During this process, the deformation of the specimen restricts the movement of the hydraulic cylinder. Consequently, the applied hydraulic energy increases to ensure that the movement of the hydraulic cylinder measured by the displacement sensor X_C matches the referenced movement X_{ref} controlled by a PLC controller and given on the basis of a predefined sheet metal bending cycle. However, the ability of the controller to fully compensate for the described causes is not given, as shown by the response error ΔX_C of the hydraulic press, which is evaluated as the difference between the measured X_C and the referenced X_{ref} displacement of the hydraulic cylinder. In addition, the pressure sensors are integrated to monitor the pressure conditions p_A, p_B in both hydraulic cylinder chambers. In addition, a force sensor is positioned between the hydraulic cylinder and the pressing plate to monitor the generated hydraulic cylinder pressing force F_C. The integrated LVDT sensor allows to foresee the opening of hydraulic valve X_V. The intensity of the bending process was controlled by applying combinations of hydraulic cylinder velocities (5 mm/s, 15 mm/s, 25 mm/s) and the width of the samples (10 mm, 20 mm, 30 mm, 40 mm). To capture the analog signals from the sensors, Beckhoff modules were used to import the data into the digital datasets. Each scenario was experimentally performed three times to confirm the repeatability of the experiments.

The friction conditions in the press guides arise due to the friction between two sliding surfaces. Stribeck friction, which is characterised by its dynamic nature, is defined by the correlation between frictional force and velocity of movement [7]. The choice of sealing and guiding technology in hydraulic cylinders is influenced by the quality and stability of the dynamic friction of hydraulic cylinders [25]. Advanced guiding and sealing systems enable a more stable friction profile over the entire range (static, hydrodynamic and mixed friction). In contrast, conventional sealing and guiding systems exhibit characteristics in which the static friction phase is significantly higher than the hydrodynamic phase. In addition, higher frictional forces occur in conventional sealing and guiding systems under the same velocity conditions. The implementation of a subsystem with three pulleys and a cable allows the emulation of different friction scenarios during the phases of the sheet metal bending cycle. By adjusting the load, the desired friction scenario is achieved and measured directly by a force gauge attached to the pressing plate, which reflects the resistance force during the movement of the hydraulic cylinder.

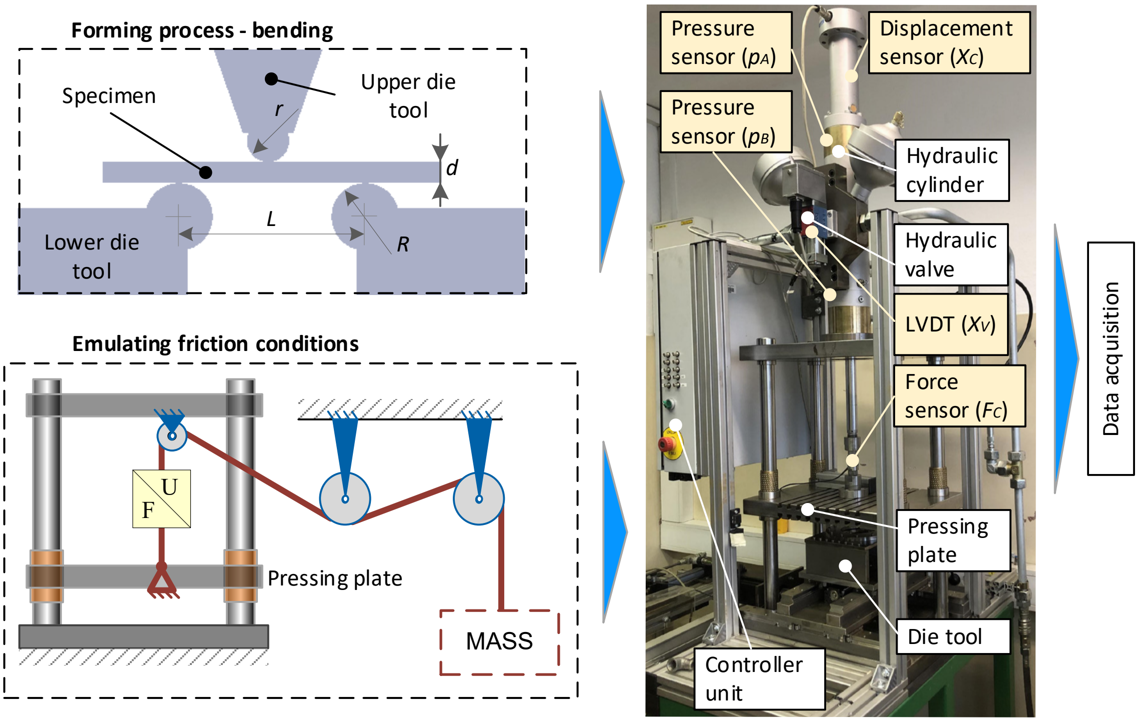

Fig. 1 Hydraulic press condition analysis

2.2 Experimental analysis

The data is visualised using variables measured in real time (sampling frequency of 100 Hz): Pressure in the upper and lower hydraulic cylinder chambers p_A, p_B, hydraulic valve opening X_V, hydraulic cylinder force F_C and hydraulic cylinder displacement X_C. The visualisation and interpretation of the raw data are the most important steps in data analytics. The extraction of relevant data intervals enables data-driven models to deliver enhanced prediction performance [26]. Furthermore, Fig. 2 shows the comparison of two main causes and their effects on the hydraulic press response, with the least intensive and the most intensive scenario shown for each cause to illustrate which measured variable offers the most information in each scenario.
The sheet metal bending cycle enhances three phases:

- Fast- forward movement (1) and fast- backward movement (4) of the hydraulic cylinder to minimise the duration of the sheet metal bending cycle and enable the production of as many parts as possible within the allotted time frame.
- In the bending phase (2), a bending force is applied to deform a sheet metal workpiece and bring it into the desired shape.
- Levelling phase (3) includes processes aimed at achieving uniformity and flatness of the sheet metal workpiece after the bending phase, with the die tool remaining stationary in its position.

The presented scenarios depict a selected sheet metal bending cycle, with the hydraulic cylinder velocity set to 50 mm/s during the movement phase (phases 1 & 4) and 15 mm/s during the bending phase (phase 2). Looking at the cause of the forming process shown in Fig. 2(a), the first scenario represents the intensity of unrestricted movement without a test specimen, while the second scenario shows the highest forming intensity, involving the placement of a specimen with a width of 40 mm is inserted into the die tool. The obvious changes occur when the die tool comes into contact with the specimen, resulting in a significant increase in the forming force F_C, which changes during the bending phase and remains semi-stationary during the levelling phase. The same tendency is observed for the variable p_A, while the variable p_B decreases as expected. During this event, the opening of the hydraulic valve X_V increases to compensate for the hydraulic press response error ΔX_C, which persists until the end of the levelling phase, when the pressing plates release the elastic tension in the sample.

The analysis of different friction scenarios shown in Fig. 2(b) indicates non-visible changes in the hydraulic cylinder displacement X_C, which can be seen by averaging the value over the phase range. The first scenario signifies more favourable friction conditions with no restraining force, while the second scenario conditions lower friction characteristics, simulated by a pulley system emulating a restraint force of 1 kN. The most noticeable changes manifest in a visible rise in the forming force during the sheet metal bending cycle and the variable p_A, accompanied by an expected decrease in the value of the variable p_B. In addition, a slight increase in the hydraulic valve opening is observed. Clearly, the influence of the friction conditions on hydraulic press response error is less significant compared to the impact observed in the case of the forming process, especially when considering maximum force F_C required to deform the sample at 15 kN.

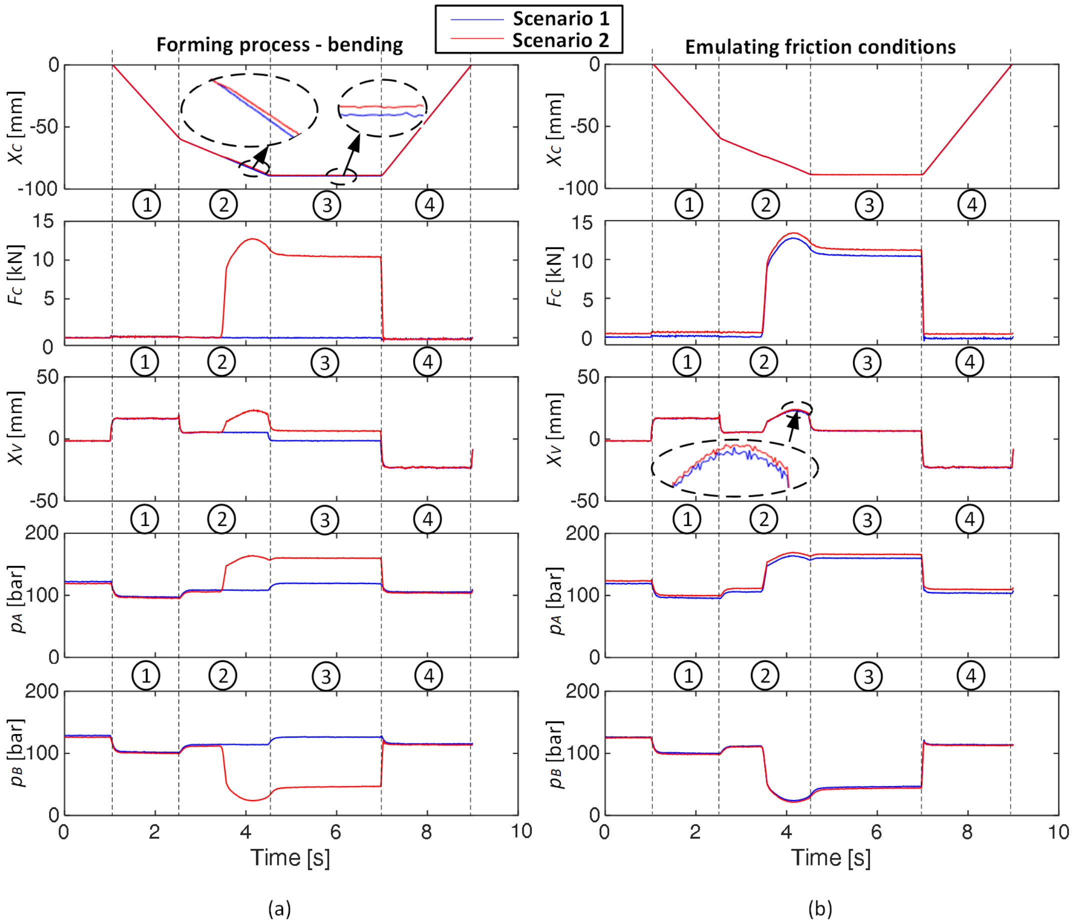

Fig. 2 Data visualization of measured variables of: (a) forming process-bending and (b) friction conditions

2.3 Data pre-processing

In regression learning, the crucial step of feature selection involves identifying and include relevant input variables, thereby enhancing the accuracy and predictive ability of the model. Moreover, regression modelling is a statistical method that reveals features that are not readily apparent in the visualisation phase, yet provides indispensable insights into data characteristics that may otherwise go unnoticed [16]. In addition, the pressure difference Δp in hydraulic cylinder chambers is assumed to be the difference between the variables p_A and p_B, providing a higher intensity of information [27]. In addition, the hydraulic cylinder velocity v_C is considered more meaningful than the measured hydraulic cylinder displacement ΔX_C, as it characterises much more dynamic conditions [28]. While the purpose of regression models is to predict the hydraulic press response error ΔX_C, the input data for regression models include the variables ΔX_C, F_C, X_V, v_C and Δp.

Industry 4.0 underlines the importance of using smart data to minimize latency and optimize the efficiency of data transmission [29]. The case study includes an investigation of 76 different conditions under which the hydraulic press was examined. However, it's crucial that the raw data for each condition includes information from different scenarios. To address this, the data must be pre-processed separately and combined into one dataset covering all scenarios so that the regression model can account for all scenarios. In addition, employing the data bundling enables the packaging of only pertinent sections into a unified dataset for each phase of the sheet metal bending cycle, as shown in Fig. 3. During the pre-processing stage, the variable sections were carefully selected to illustrate the semi-stable operating condition of the hydraulic press over a time scale. Furthermore, the extracted sections were selected considering the visualization aspects described in chapter 2.2. In addition, a feature scaling was performed to standardize the importance of the five variables and to prevent any single variable from being prioritized [30]. The scaling was performed to normalize the y-axis by setting it in the range from 0 to 1. Finally, the extracted datasets representing different scenarios were partitioned into training and test datasets to evaluate the regression models using known metrics, including Root Mean Square Error (RMSE), coefficient of determination R^2, Mean Squared Error (MSE) and Mean Average Error (MAE). Furthermore, training datasets are used during the model training phase to determine the required training time for different regression model approaches. In addition, test datasets were used to evaluate the prediction accuracy of each model after the training phase, considering the testing time required to perform the predictions.

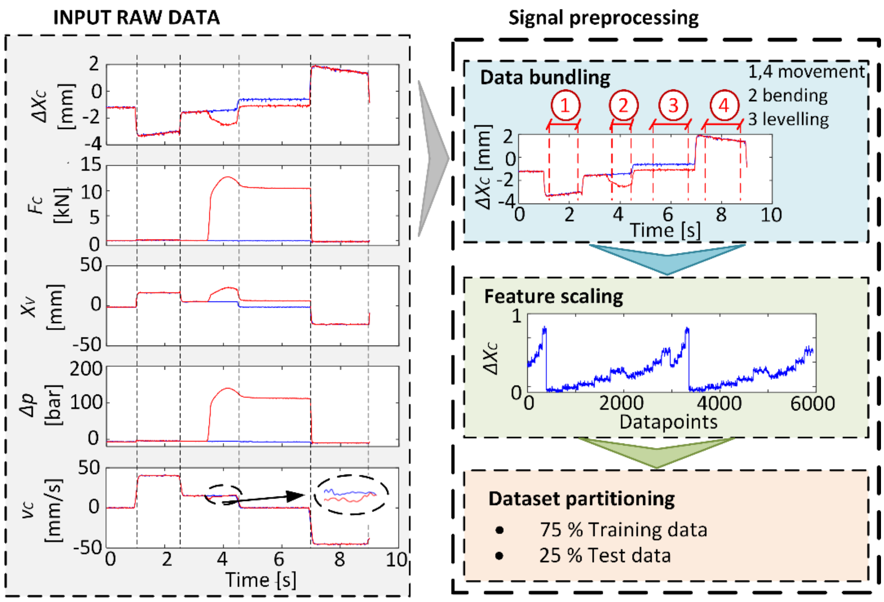

Fig. 3 Smart data forming

3. Regression models

The availability of regression modelling methods offers a wide range of possibilities to achieve semi-mirrored data-driven model for the chosen research problem. Grid search and random search are two commonly employed methods for optimizing hyperparameters. While grid search exhaustively explores all possible parameter combinations, it incurs high computational costs, especially when a large number of parameters are involved [18]. In contrast, the random search examines potential parameter combinations randomly within fixed ranges and offers a significant reduction in computational cost. In our study, both approaches are used for hyperparameter optimization. Grid search is applied to LR, GPR and SVR, while random search is applied to DT and NN depending on the number of hyperparameters involved. The MATLAB application Regression Learner was used to train and test the regression models.

Five-fold cross-validation was used to determine the hyperparameters that provide the best generalization performance for our models. In each iteration, a specific subset was selected for testing, while the remaining subsets were used for training. The investigated hyperparameters for the models are summarized for each model in its subsection.

3.1 Linear regression

Linear regression (LR) uses an assumption function described in Eq. 1 to model and capture the relationships between independent variables x, where β is the coefficient vector, n is the number of samples, and y is the target value [12]. In our approach, we include interaction terms in the linear regression model; nevertheless, we also use the classical LR approach. This decision is based on previous studies that indicate a varying degree of influence of different independent variables on the dependent variable [24]. Including interaction terms allows the model to capture and account for the joint effects of these variables, taking into account the potential dependencies and relationships identified in previous research [12].

The objective of minimizing the objective function is to reduce the sum of squared differences between the observed values and the corresponding predicted values, as described in Eq. 2 by the least squares method.

$$y(x) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_2 (x_1 \cdot x_2) + \cdots + \beta_n x_n \tag{1}$$

$$J(\beta) = \frac{1}{2n} \sum_{i=1}^{n} (y(x_i) - y_i)^2 \tag{2}$$

3.2 Decision trees

Decision trees assume nonlinear patterns and relative relationships in the data and perform feature selection for the most accurate prediction [13]. In addition, DT have simple extraction rules, are very accurate and provide good interpretability of the models. In our approach, the search for the optimal parameters for DT involves varying the leaf size (4, 12 and 36).

3.3 Support vector regression

The SVR is an extension of support vector machine developed for single-output regression [14]. In addition, the SVR determines a linear dependence between the independent variable x and the dependent variable y, as shown in Eq. 3, where w is the weight vector and b is the intercept.

$$y = wx + b \tag{3}$$

The goal of objective function in SVR is to find the weight vector w and the intercept b such that configures the least deviation in between the predicted and actual values as expressed in Eq. 4. Here, a regularization parameter C balances the trade-off between w and the slack variables ξ, ξ^*, under the conditions expressed in Eq. 5. Fig. 4 describes the linear kernel function approximation considering the data accuracy of ε.

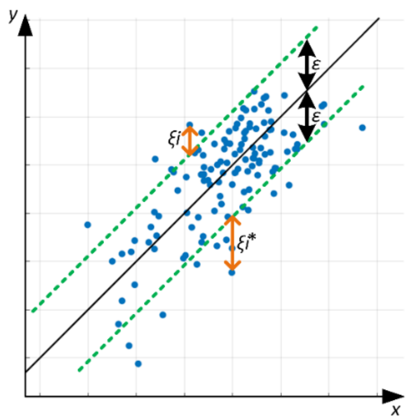

Fig. 4 Linear kernel function for SVR

$$J(w,b,\xi,\xi^*) = \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n}(\xi_i + \xi_i^*) \tag{4}$$

$$\begin{cases} y_i - wx_i - b \leq \varepsilon + \xi_i \\ wx_i + b - y_i \leq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* > 0 \end{cases} \tag{5}$$

The decision function in SVR with a linear kernel is determined in Eq. 6.

$$y = \sum_{i=1}^{n}(\alpha_i - \alpha_i^*)\,K(x,x_i) + b \tag{6}$$

Enhancing the accuracy of nonlinear predictions, we consider a number of kernel options described in Eq. 6 through 10, including linear, quadratic K_q, cubic K_c, exponential K_e, and squared exponential K_{se}, where c is a constant term. These kernels are systematically evaluated using a random search approach to uncover the most appropriate SVR model.

$$K_q = (x + c)^2 \tag{7}$$

$$K_c = (x + c)^3 \tag{8}$$

$$K_e = \exp\left(-\frac{\|x_i - x_j\|}{l}\right) \tag{9}$$

$$K_{se} = \exp\left(-\frac{\|x_i - x_j\|^2}{2l^2}\right) \tag{10}$$

3.4 Gaussian process regression

Due to the non-parametric Bayesian approach to regression problems, the GPR offers a high degree of flexibility for the targeted prediction of complex relationships in the data [16]. It assumes a Gaussian data distribution, where the model hyperparameters, i.e. the variance of the distribution σ_f and the length parameter l, are determined by the mean function $m(x)$ and the kernel K, as shown in Eq. 11 [31]. The purpose of the mean function is to represent the expected trend of the decision function y, which is assumed to be a zero-mean function in our case [24]. In addition, the purpose of the kernel function is to assign importance rates to the training datasets and to fit the model.

$$y \sim GP\,(m(x), K) \tag{11}$$

To extract the most accurate GPR model, a σ_f of 0.001-0.100 was varied. In the training process, Automatic Relevance Determination (ARD) allows the model to automatically determine the relevance of each predictor by adjusting the length parameter l, considering different types of kernels as shown in Eq. 6 to 10.

3.5 Neural network

Neural networks consist of neurons organized in layers that determine the connections between neurons through a series of weights and biases [22]. The input data is processed to provide an output of each neuron, which allows flexibility in determining the number of neurons in each layer and the number of layers in the neural network to best fit the data. Neural networks are composed of an input layer, one or more hidden layers and an output layer of artificial neurons, as shown in Fig. 5. Artificial Neural Network (ANN) is a broad term that encompasses NN models of varying complexity, while Deep Neural Network (DNN) specifically refers to networks with multiple hidden layers that emphasize the depth of the architecture, which is usually associated with an increased ability to learn complicated representations from data [23]. Models built in our case with NN include five neurons in the input layer, since the available predictors are five. The output layer consists of one neuron, since only one output estimate is predicted.

In the configuration of the neural network, we varied the number of neurons in the hidden layer between 10, 25 and 100, while the number of layers ranged from one to three. ReLU was chosen as the activation function, with a maximum of 1000 iterations and a regularization strength of 0.

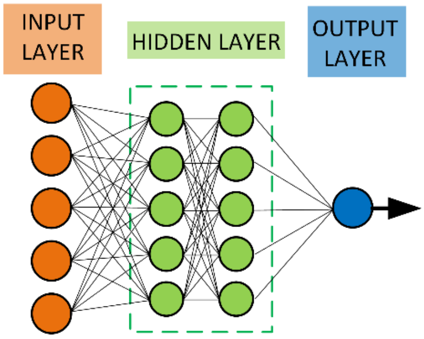

Fig. 5 Neural network structure

4. Results and discussion

4.1 Characteristics of fitted regression models

The performance of regression models of different types depends on model complexity, data quality and quantity, feature selection, assumption violations, pre-processing stage and hyperparameter selection [18]. The results shown in Table 1 represent the best-fit regression models with the most optimal selection of hyperparameters for the given case study and the range of regression model trained and test fitted in the MATLAB environment. Considering that the research problem focuses on three phases of the sheet metal bending cycle, i.e. bending, levelling movement, the complexity of the regression model differs due to the nature of the data patterns, the presence of white noise and the number of conditions summarised in each phase.

Regression modelling for LR yields superior results when interaction terms are included across all phases of the sheet metal bending cycle than the method without interaction terms. In addition, the DT type of regression modelling shows the importance of selecting hyperparameter variation. In the bending and levelling phase, the necessary depth of the DT is significantly higher than in the case of the movement phase. In addition, the medium size tree is mostly appropriate for the bending phase, while the coarse tree size is the most appropriate for the levelling and movement phases, considering the number of leaves and the size of the parents. The results in the table for the SVR regression modelling type show that the squared exponential kernel function is most effective in representing complex relationships and patterns among the input data. In addition, the hyperparameters C and ε reflect the trade-off between the accurate fit of the training data and the margin of tolerance, which is comparable between phases. Moreover, selected hyperparameters with low values indicate an easy fit of the model and a relatively narrow range, emphasising the minimization of errors in the predictions. In addition, the SVR regression model shows that less complexity is required in the case of the bending phase and the margin of error is even narrower than in the SVR regression models for the levelling and movement phases. Moreover, the results of the GPR regression models show that the Exponential kernel function consistently provides the most optimal fit among the different kernels available. In the levelling phase, it is evident that a higher standard deviation of noise is required as the data has a higher amount of white noise compared to the movement and bending phase. For the application of regression modelling using NN, it has been shown that a single layer is sufficient for the levelling and bending phases. However, for optimal performance in the levelling phase, it is best to use three layers with a layer size of 10.

Table 1 Best fitted regression models summary

Regression model	Hyperparameters and characteristics	Bending phase	Levelling phase	Movement phase
Linear regression (LR)	Interactions terms	ON	ON	ON
	Robust option	OFF	OFF	OFF
Decision Trees (DT)	Max. depth	410	390	162
	Min. leaf size	12	36	36
	Min. Parent size	24	72	72
	Max. Number Splits	4454	9067	8666
	Split Criterion	MSE	MSE	MSE
	Prune	ON (RMSE)	ON (RMSE)	ON (RMSE)
	Surrogate	OFF	OFF	OFF
Support Vector Regression (SVR)	Cost C	0.1848	0.2708	0.2768
	Epsilon ε	0.0185	0.0271	0.0277
	Kernel type	K_{SE}	K_{SE}	K_{SE}
	Kernel Scale	0.5	0.5	2
	Iteration limit	10^6	10^6	10^6
Gaussian Process Regression (GPR)	Kernel type	K_E	K_E	K_E
	Standard deviation of noise σ_n	0.0100	0.0438	0.0038
	Quasi-Newton optimization	ON	ON	ON
Neural Network (NN)	Number of layers	1	1	3
	Layer size	100	100	10, 10, 10
	Activation	ReLu	ReLu	ReLu
	Iteration Limit	1000	1000	1000
	Regularization strength	0	0	0

4.2 Regression model performance evaluation

The overall performance of the regression models was assessed using the validation and test results to confirm that the data-driven models do not overfit. The results for the bending phase are presented in Table 2 and show significant performance accuracy in the validation stage by the R^2 metric, which is at a value between 0.9945 and 0.9961. In addition, the complexity of the regression modelling affects the performance of the regression models. This is particularly the case for less complex regression modelling approaches such as LR and DT. In such examples, the values of the LR regression model for the metrics RMSE, R^2, MSE, MAE are 0.01364, 0.9945, 0.000186, 0.009998 in the validation and 0.01368, 0.9941, 0.000187, 0.009836 in the test. As the complexity of the regression modelling increases, the performance of the regression model increases, which is evident in the regression models GPR and NN by higher evaluation metrics RMSE, R^2, MSE, MAE. In comparison, the best fitted regression model in the case of the bending phase is the GPR model with metric values RMSE, R^2, MSE, MAE at validation 0.01144, 0.9961, 0.000131, 0.007810 and at test 0.01150, 0.9958, 0.000132, 0.007768. Although more complex regression models, e.g., GPR and NN, perform better, less complex regression models, e.g., LR and DT, accurately predict the desired response error of the hydraulic press in different scenarios. In addition, the performances of the regression models built by different approaches provide solid results in the test stage, where the coefficient of determination is between 0.9941 and 0.9958.

Table 2 Regression model performance evaluation for bending phase

Model	Validation				Test			
	RMSE	R^2	MSE	MAE	RMSE	R^2	MSE	MAE
LR	0.01364	0.9945	0.000186	0.009998	0.01368	0.9941	0.000187	0.009836
DT	0.01357	0.9946	0.000184	0.009316	0.01265	0.9950	0.000160	0.008854
SVR	0.01252	0.9954	0.000157	0.009066	0.01238	0.9952	0.000153	0.008754
GPR	0.01144	0.9961	0.000131	0.007810	0.01150	0.9958	0.000132	0.007768
NN	0.01168	0.9960	0.000136	0.008105	0.01156	0.9958	0.000134	0.007890

The validation and test results for the levelling phase are shown in Table 3. From the metrics, the best regression model in validation is the DT model, which has the lowest RMSE, MSE and MAE at 0.04393, 0.001930 and 0.030141, but the highest R^2 at 0.9618. In addition, the other fitted regression models considering the metrics of RMSE, R^2, MSE, MAE are in the order of GPR, SVR, NN, LR. In addition, the test results show that the same sequence of regression models enables the best regression model performance. Moreover, the overall accuracy of the different regression models is solid, as R^2 is between 0.9517 and 0.9618 for the validation results and 0.9483 and 0.9581 for the test results. Compared to the bending phase, the performance of the regression models is lower due to the white noise contained in the experimental data.

In the movement phase, the best-fitting regression models, i.e. DT, GPR and NN, have a robust coefficient of determination R^2 of 0.9995 for both validation and test results, as shown in Table 4. In addition, DT is shown to be the best performing model with the lowest values for RMSE, MSE and MAE. The next best performing regression model for the movement phase is NN, with GPR being similarly accurate to NN in terms of the RMSE, MSE and MAE metrics. The lowest performing regression model is SVR with an R^2 value of 0.9986 in validation and 0.9987 in test. In addition, the LR model is more accurate than SVR in prediction. Overall, the prediction accuracy of the regression models is higher in the movement phase than in other phases of the sheet metal bending results, which is evident by an R^2 value of more than 0.999 in the validation and test. Overall, regression models are built accurately and are not overfitting trough all phases of the sheet metal bending cycle, i.e. bending, levelling and movement, considering that the RMSE, R^2, MSE and MAE metrics show similar values between validation and test results.

Table 3 Regression model performance evaluation for levelling phase

Model	Validation				Test			
	RMSE	R^2	MSE	MAE	RMSE	R^2	MSE	MAE
LR	0.04944	0.9517	0.002445	0.036453	0.05070	0.9483	0.002571	0.037562
DT	0.04393	0.9618	0.001930	0.030141	0.04568	0.9581	0.002086	0.031166
SVR	0.04770	0.9550	0.002275	0.034339	0.04801	0.9537	0.002305	0.034739
GPR	0.04628	0.9576	0.002142	0.033074	0.04680	0.9560	0.002190	0.033308
NN	0.04829	0.9539	0.002332	0.035171	0.04830	0.9531	0.002333	0.035264

Table 4 Regression model performance evaluation for movement phase

Model	Validation				Test			
	RMSE	R^2	MSE	MAE	RMSE	R^2	MSE	MAE
LR	0.08885	0.9991	0.007894	0.068134	0.08894	0.9991	0.007909	0.067491
DT	0.06629	0.9995	0.004395	0.048779	0.06658	0.9995	0.004433	0.048933
SVR	0.11099	0.9986	0.012318	0.089846	0.10939	0.9987	0.011966	0.089033
GPR	0.06781	0.9995	0.004598	0.050394	0.06801	0.9995	0.004626	0.050394
NN	0.06719	0.9995	0.004515	0.050851	0.06709	0.9995	0.004501	0.049864

4.3 Training and prediction rate analysis

The comparison of the training and test times in between the regression models LR, DT, SVR, GPR and NN is shown in Fig. 6. In terms of training and prediction time, the regression models LR, DT and SVR show better results for all phases of the sheet metal bending cycle, i.e. bending, levelling movement, compared to more complex regression models such as GPR and NN.

Examining the bending phase shown in Fig. 6(a), it is evident that LR, DT and SVR models exhibit short training times below 5 s and short prediction times, which are less than 8.5 ms. The best training and prediction rate were achieved by the DT model with a training time of 3.8 s and a prediction time of 3.4 ms. In contrast, the training time of the GPR and NN models is 190.6 s and 183.7 s, which corresponds to a 50 times higher rate compared to the less complex DT model. In addition, the GPR model requires 59.9 ms to predict the response error of the hydraulic press in the bending phase, compared to the NN model prediction time of 3.9 ms. Overall, the prediction rate of the NN model is comparable to other low-complexity regression models, as evidenced by the test time of LR, DT and SVR of up to 8.4 ms.

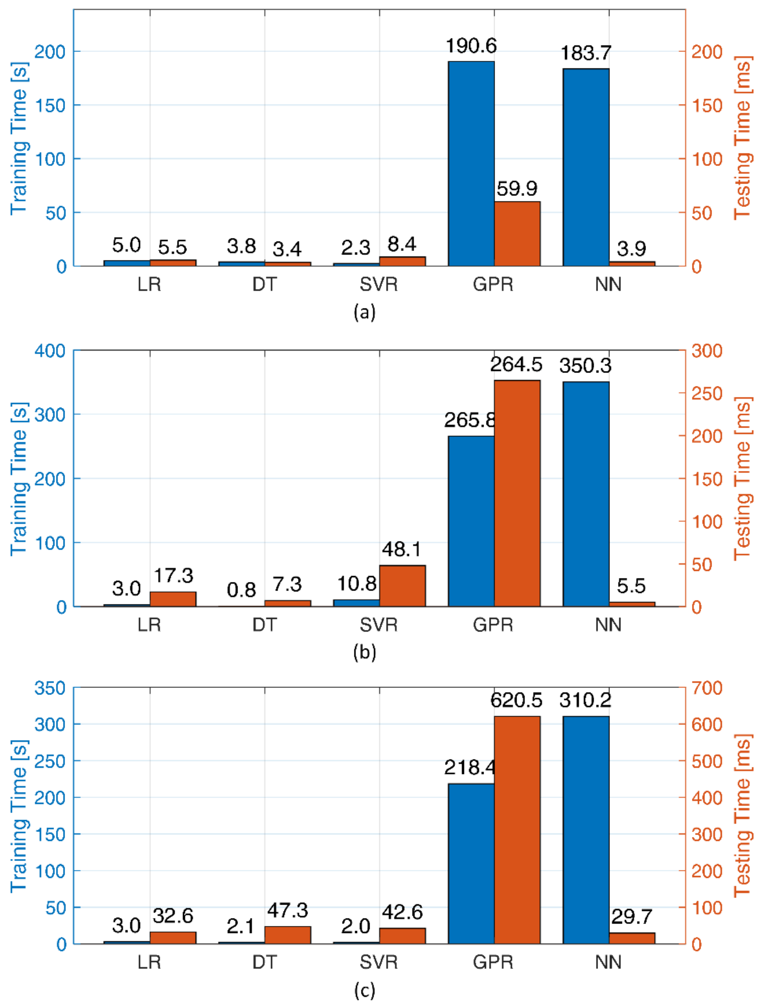

Fig. 6 Regression model training and testing time comparison for: (a) bending phase, (b) levelling phase, (c) movement phases

In the case of the levelling phase shown in Fig. 6(b), the training rate for the LR, DT and SVR models is up to 437 times better than for the GPR and NN models. In addition, the training time of the DT model is solid at 0.8 s for the entire training dataset. The training time of the NN model is the longest at 350.3 s, but the prediction time is the shortest at 5.5 ms compared to the other models. In comparison, the SVR model requires 6 times more time to predict the outcome than the DT model. The GPR model has the highest prediction rate with a prediction time of 264.5 ms. Examining the movement phase shown in Fig. 6(c), the LR, DT and SVR models require a training time of up to 3 s, which is on average better than in the bending and levelling phase. However, the LR, DT and SVR models require higher prediction times on average compared to the bending and levelling phase. In the movement phase, datasets size is higher compared to the bending and levelling phase, but the required complexity of the regression model is lower. In addition, due to the larger datasets, the number of predictions is also higher, which leads to higher prediction times. In terms of prediction time, the NN model has the lowest prediction time of 29.7 ms. In addition, the GPR model requires the longest prediction time of 620.5 ms, which is 20 times higher compared to the NN model.

4.4 A general framework for enhancing production systems using machine learning

The study also aims to provide general guidelines for transforming of conventional production systems to a smart production system using ML techniques. The main steps from 1 to 4 can be used to define the most suitable machine learning strategy and approach for particular production use case such as robotic assembly, CNC machining, etc. (Fig. 7).

Step 1: The effectiveness of these methods depends on the availability and quality of the data required for model training (Step 1, Data collection). Focusing on our use case, the results of the hydraulic press show the importance of tailored approaches to improve process control and the performance of the different data samples from production processes in terms of pattern, complexity, presence of Gaussian noise, dimensionality and transition types. In general, the most important is to design experiments and data collection approaches, focusing on selecting appropriate sensors and determining their location to extract the relevant information from the real production system and ensure the validity of the information collected. The sensors serve as the primary data source for monitoring specific characteristics and requirements relevant to the research. Furthermore, in step 4, the input variables for ML modelling are determined, which influence the predictive accuracy of the outcome variables and summarize the specific characteristics and requirements of the respective research.

Step 2: Here, the data analysis is performed by visualizing the various conditions of the production system under investigation. Learning and testing datasets should be extracted from different condition scenarios at intervals to observe the most relevant sections of the measured data for predicting outcomes of ML model.

Step 3: This step represents the signal pre-processing, which involves the condition recognition of production system. In the case of hydraulic presses condition recognition, the most important task is data pre-processing that involves bundling and scaling. In addition, when observing other production processes their characteristics are different, thereby other pre-processing methods, such as logarithmic transformation may be more appropriate to consider.

Step 4: Finally, when selecting a machine learning strategy, different ML approaches should be tested to determine the most appropriate one for the application in terms of given criteria and objective function. For hydraulic presses, the edge computing decision-making algorithm integrated into the control system is expected to be fast and accurate in order to improve the performance of the hydraulic press in real-time. The prediction rate significantly impacts the ability of the hydraulic press control system to adjust responses autonomously, increasing production reliability and reducing product uncertainty.

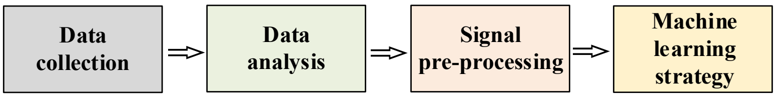

Fig. 7 Guidelines for accurate machine learning integration into production processes

5. Conclusion

This research addresses the trade-off between model complexity, prediction accuracy, and computational efficiency. In real-time applications, rapid response times are crucial for effective reactions. Therefore, the efficiency of a decision-making system that reacts to recognised conditions in the production processes depends on the prediction rate and the accuracy of the ML model. In addition, the adaptability of ML models to newly introduced conditions are determined by the learning rate. While more complex models such as GPR and NN deliver higher prediction accuracy, the associated training and prediction times need to be carefully considered, especially in different phases of the sheet metal bending cycle i.e. forming, levelling and movement. The varying performance of LR, DT, and SVR models showcases the importance of selecting models tailored to the specific requirements and constraints of the application. In conclusion, the metric R^2 yielded robust results in validation and test performance accuracy in each phase of the sheet metal bending cycle: bending 0.995, levelling 0.950, and movement 0.999. The findings suggest that more complex regression models exhibit superior performance, but they are associated with a higher training time, which is up to 437 times higher than for less complex models such as LR, DT and SVT. Notably, GPR requires even longer prediction times compared to other ML models. The difference in prediction time could be attributed to the explicit representation and computational intricacies involved in GPR during predictions, whereas NN often benefit from more implicit representations and parallelization, leading to faster predictions. Furthermore, the memory requirement with models NN is much lower compared to GPR, as GPR needs to store information for the entire training dataset, including covariance matrices. Models LR, DT, and SVR can be treated as lower-performance models, however evidently need shorter training and prediction times for the given case study. Therefore, the optimal ML models for integration into the decision-making system are either the less complex LR model, offering higher training and prediction rate, or either more intricate NN, which requires a higher training rate. However, despite higher training rates, NN showcases superior prediction accuracy and prediction rate comparable to other less complex ML models.

In summary, the design of the control system for the hydraulic press allows the consideration of predetermined operating conditions, however the occurrence of new conditions would require the optimization of machine learning hyperparameters. The methodology presented in this study provides a valuable reference point for the selection of modelling and solution methods suitable for the diagnosis of autonomous production process conditions. Given the complexity of production processes and the desired prediction results, an intelligent combination of different solution methods can effectively address these challenges. By applying multiple machine learning methods such as LR, GPR, NN and SVR, this study presents the most suitable approach that enables accurate and rapid prediction of hydraulic press response.

Acknowledgement

The work was carried out in the framework of the Slovenian Research and Innovation Agency (research program funding No. P2-0248, research project No. J2-4470 and research program for young researchers No. 53512) funded by the Republic of Slovenia – Ministry of Education, Science and Sport. The work was carried out in the framework of research project STAGE – Sustainable Transition to the Agile and Green Enterprise funded under the Grant agreement ID: 101058693.

References

[1] Li, J., Lu, Y., He, F., Miao, L. (2023). Motion control of hydraulic actuators with nonlinear friction compensation: Applied to variable valve systems of diesel engine, *ISA Transactions*, Vol. 137, 561-573, doi: 10.1016/j.isatra.2023.01.037.

[2] Stavropoulos, G., Violos, J., Tsanakas, S., Leivadeas, A. (2023). Enabling artificial intelligent virtual sensors in an IoT environment, *Sensors*, Vol. 23, No. 3, Article No. 1328, doi: 10.3390/s23031328.

[3] Hariri-Ardebili, M.A., Mahdavi, G., Nuss, L.K., Lall, U. (2023). The role of artificial intelligence and digital technologies in dam engineering: Narrative review and outlook, *Engineering Applications of Artificial Intelligence*, Vol. 126, Part A, Article No. 106813, doi: 10.1016/j.engappai.2023.106813.

[4] Dias, I., Ruan, L., Ranaweera, C., Wong, E. (2023). From 5G to beyond: Passive optical network and multi-access edge computing integration for latency-sensitive applications, *Optical Fiber Technology*, Vol. 75, Article No. 103191, doi: 10.1016/j.yofte.2022.103191.

[5] Chen, X., Liu, H., Nikitas, N. (2023). Internal pump leakage detection of the hydraulic systems with highly incomplete flow data, *Advanced Engineering Informatics*, Vol. 56, Article No. 101974, doi: 10.1016/j.aei.2023.101974.

[6] Kang, F., Wu, Y., Li, J., Li, H. (2021). Dynamic parameter inverse analysis of concrete dams based on Jaya algorithm with Gaussian processes surrogate model, *Advanced Engineering Informatics*, Vol. 49, Article No. 101348, doi: 10.1016/j.aei.2021.101348.

[7] Pan, Q., Li, Y., Huang, M. (2018). Control-oriented friction modeling of hydraulic actuators based on hysteretic nonlinearity of lubricant film, *Mechatronics*, Vol. 53, 72-84, doi: 10.1016/j.mechatronics.2018.05.003.

[8] Feng, H., Qiao, W., Yin, C., Yu, H., Cao, D. (2019). Identification and compensation of non-linear friction for a electro-hydraulic system, *Mechanism and Machine Theory*, Vol. 141, 1-13, doi: 10.1016/j.mechmachtheory. 2019.07.004.

[9] Karanović, V., Jocanović, M., Baloš, S., Knežević, D., Mačužić, I. (2019). Impact of contaminated fluid on the working performances of hydraulic directional control valve, *Journal of Mechanical Engineering*, Vol. 65, No. 3, 139-147, doi: 10.5545/sv-jme.2018.5856.

[10] Zhu, Y., Li, G., Tang, S., Wang, R., Su, H., Wang, C. (2022). Acoustic signal-based fault detection of hydraulic piston pump using a particle swarm optimization enhancement CNN, *Applied Acoustics*, Vol. 192, Article No. 108718, doi: 10.1016/j.apacoust.2022.108718.

[11] Ji, X., Ren, Y., Tang, H., Shi, C., Xiang, J. (2020). An intelligent fault diagnosis approach based on Dempster-Shafer theory for hydraulic valves, *Measurement,* Vol. 165, Article No. 108129, doi: 10.1016/j.measurement.2020.108129.

[12] Ahmadizadeh, C., Menon, C. (2019). Investigation of regression methods for reduction of errors caused by bending of FSR-based pressure sensing systems used for prosthetic applications, *Sensors*, Vol. 19, No. 24, Article No. 5519, doi: 10.3390/s19245519.

[13] Shi, M., Hu, W., Li, M., Zhang, J., Song, X., Sun, W. (2023). Ensemble regression based on polynomial regression-based decision tree and its application in the in-situ data of tunnel boring machine, *Mechanical Systems and Signal Processing*, Vol. 188, Article No. 110022, doi: 10.1016/j.ymssp.2022.110022.

[14] Tang, X., Wu, C., Xu, X. (2022). Learning-based nonlinear model predictive controller for hydraulic cylinder control of ship steering system, *Journal of Marine Science and Engineering*, Vol. 10, No. 12, Article No. 2033, doi: 10.3390/jmse10122033.

[15] Tang, S., Zhu, Y., Yuan, S. (2022). An adaptive deep learning model towards fault diagnosis of hydraulic piston pump using pressure signal, *Engineering Failure Analysis*, Vol. 138, Article No. 106300, doi: 10.1016/j.engfailanal.2022.106300.

[16] Kang, F., Wu, Y., Ma, J., Li, J. (2023). Structural identification of super high arch dams using Gaussian process regression with improved salp swarm algorithm, *Engineering Structures*, Vol. 286, Article No. 116150, doi: 10.1016/j.engstruct.2023.116150.

[17] Sami, I., Ullah, S., Ullah, N., Ro, J.-S. (2021). Sensorless fractional order composite sliding mode control design for wind generation system, *ISA Transactions*, Vol. 111, 275-289, doi: 10.1016/j.isatra.2020.11.001.

[18] Su, W., Ren, W., Sun, H., Liu, C., Lu, X., Hua, Y, Wei, H., Han, J. (2022). Data-based flow rate prediction models for independent metering hydraulic valve, *Energies*, Vol. 15, No. 20, Article No. 7699, doi: 10.3390/en15207699.

[19] Lei, Y., Jiang, W., Jiang, A., Zhu, Y., Niu, H., Zhang, S. (2019). Fault diagnosis method for hydraulic directional valves integrating PCA and XGBoost, *Processes*, Vol. 7, No. 9, Article No. 589, doi: 10.3390/pr7090589.

[20] Zhong, Q., Xu, E., Shi, Y., Jia, T., Ren, Y., Yang, H., Li, Y. (2023). Fault diagnosis of the hydraulic valve using a novel semi-supervised learning method based on multi-sensor information fusion, *Mechanical Systems and Signal Processing*, Vol. 189, Article No. 110093, doi: 10.1016/j.ymssp.2022.110093.

[21] Liu, Y., Li, W., Lin, S., Zhou, X., Ge, Y. (2023). Hydraulic system fault diagnosis of the chain jacks based on multi-source data fusion, *Measurement*, Vol. 217, Article No. 113116, doi: 10.1016/j.measurement.2023.113116.

[22] Shweta, R., Sivagnanam, S., Kumar, K.A. (2023). IoT-based deep learning neural network (DLNN) algorithm for voltage stability control and monitoring of solar power generation, *Advances in Production Engineering & Management*, Vol. 18, No. 4, 447-461, doi: 10.14743/apem2023.4.484.

[23] Guo, Y., Zeng, Y., Fu, L., Chen, X. (2019). Modeling and experimental study for online measurement of hydraulic cylinder micro leakage based on convolutional neural network, *Sensors*, Vol. 19, No. 9, Article No. 2159, doi: 10.3390/s19092159.

[24] Jankovič, D., Šimic, M., Herakovič, N. (2024). A data-driven simulation and Gaussian process regression model for hydraulic press condition diagnosis, *Advanced Engineering Informatics*, Vol. 59, Article No. 102276, doi: 10.1016/j.aei.2023.102276.

[25] Ma, K., Wang, J., Gu, L. (2018). Experimental study on friction of hydraulic cylinder in different sealing systems, *MATEC Web of Conferences*, Vol. 153, Article No. 06012, doi: 10.1051/matecconf/201815306012.

[26] Xin, C., Motz, T., Fuhl, W., Hartel, A., Kasneci, E. (2023). Deep learning-based position detection for hydraulic cylinders using scattering parameters, *Expert Systems with Applications*, Vol. 232, Article No. 120892, doi: 10.1016/j.eswa.2023.120892.

[27] Siwulski, T. (2023). Comparative study of the influence of the system architecture on the accuracy of hydraulic cylinder working movements, *Applied Sciences*, Vol. 13, No. 3, Article No. 1594, doi: 10.3390/app13031594.

[28] Zhou, R., Meng, L., Yuan, X., Qiao, Z. (2022). Research and experimental analysis of hydraulic cylinder position control mechanism based on pressure detection, *Machines*, Vol. 10, No. 1, Article No. 1, doi: 10.3390 /machines 10010001.

[29] Resman, M., Pipan, M., Šimic, M., Herakovič, N. (2019). A new architecture model for smart manufacturing: A performance analysis and comparison with the RAMI 4.0 reference model, *Advances in Production Engineering & Management*, Vol. 14, No. 2, 153-165, doi: 10.14743/apem2019.2.318.

[30] Makansi, F., Schmitz, K. (2022). Data-driven condition monitoring of a hydraulic press using supervised learning and neural networks, *Energies*, Vol. 15, No. 17, Article No. 6217, doi: 10.3390/en15176217.

[31] Dudek, A., Baranowski, J. (2022). Gaussian processes for signal processing and representation in control engineering, *Applied Sciences*, Vol. 12, No. 10, Article No. 4946, doi: 10.3390/app12104946.