20th IMEKO TC10 Conference
28-29 September 2026 | Lisbon | Portugal

20th IMEKO TC10 Conference
"Measurement for Diagnostics, Optimization and Control to Support Competitiveness and Innovation"
Lisbon, Portugal, September 28–29, 2026

Feature Selection for EDM Process Optimization in Manufacturing

Anh Tuan Hoang[1,2], András Szépligeti[3], András Németh[3], Zsolt János Viharos[4,1]

[1]HUN-REN Institute for Computer Science and Control, Budapest, Hungary
[2]Doctoral School of Informatics, Eötvös Loránd University, Budapest, Hungary
[3]AQ Anton Kft., Zalaegerszeg, Hungary
[4]Faculty of Economics and Business, John von Neumann University, Kecskemét, Hungary

***Abstract* – Electrical Discharge Machining (EDM) is a widely applied manufacturing process where performance is strongly influenced by multiple interacting technological parameters. This study presents a data-driven approach for analysing and reducing the dimensionality of EDM process parameters using feature selection methods. An industrial dataset containing six input parameters and time-related outputs across three machining stages is investigated. Two analysis tasks are defined: classification of machining regimes into high-speed and low-speed categories, and refinement within the high-speed regime. Adaptive Hybrid Feature Selection (AHFS) and Model Input-Output Configuration based Search (MICS-EFS) methods are applied to identify influential parameters and optimal feature subsets. The results show that high classification accuracy of up to 96% can be achieved for regime separation, while the high-speed regime analysis remains more challenging, with 75–85% accuracy. An optimal number of features is identified for most cases, although full parameter sets are required in certain scenarios. In addition, the reduced parameter space enables simpler process control with fewer variables to be optimally tuned and significantly decreases the number of required experiments in design of experiments (DoE) for process optimisation. The findings reveal both globally dominant and stage-dependent parameters, supporting efficient process optimisation and highlighting the benefits of a two-stage analytical framework.**

Keywords* – *electrical discharge machining (EDM), feature selection, process optimization, manufacturing, machine learning, responsible consumption and production

I. INTRODUCTION

Electrical Discharge Machining (EDM) [1–2] is a well-established non-conventional manufacturing technology widely applied in the machining of hard-to-cut materials, particularly in the aerospace and energy sectors [3]. EDM plays a critical role in modern manufacturing, where high-aspect-ratio features must be produced with high precision and reliability [4–5]. These components are typically made from advanced superalloys, whose mechanical and thermal properties make conventional machining processes inefficient or impractical. Consequently, EDM has become a key enabling technology in this domain.

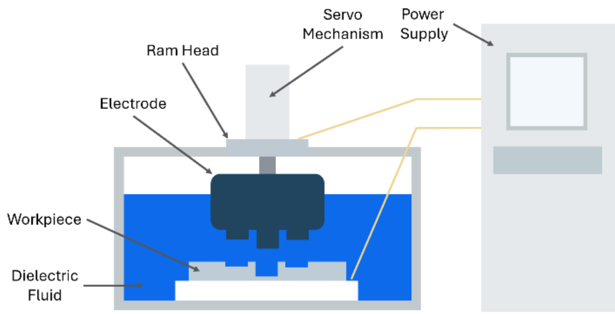

Fig. 1. EDM machine

The EDM process is based on controlled electrical discharges occurring between a tool electrode and a conductive workpiece, submerged in a dielectric fluid. Material removal is achieved through localized melting and vaporization caused by high-frequency spark discharges. The performance of the process is strongly influenced by a set of technological parameters [6], including discharge current, pulse-on time, pulse-off time, gap voltage, flushing conditions, and servo control settings. These parameters jointly determine key output characteristics such as machining time and speed, dimensional accuracy, and tool wear.

In this work, a dataset derived from controlled EDM experiments in manufacturing is analysed. The dataset consists of six key technological input parameters, along with time-related outputs measured across three machining stages: setup, processing, and completion. The primary focus of this study is on machining time as a performance indicator, while tool wear is excluded from detailed analysis due to its relatively low economic impact in the given context. The experimental data is structured to enable two levels of analysis: (i) classification of machining instances into high-speed and low-speed regimes, and (ii) refined exploration within the high-speed regime to identify parameter configurations associated with optimal performance. This two-stage analytical framework reflects practical industrial decision-making processes, where both coarse and fine-grained optimization are required.

The present study contributes to the field by providing a structured methodology for analysing and control EDM processes in an industrially relevant setting, with particular emphasis on parameter importance, its effects, and data-driven process optimization.

II. RELATED RESULTS IN THE LITERATURE

EDM process optimization has been extensively studied due to the strong sensitivity of machining performance to technological parameters. Singh et al. [6] showed that performance is typically evaluated using material removal rate, tool wear rate, relative wear ratio, and surface roughness, and is mainly influenced by parameters such as discharge current, pulse-on/off time, arc gap, duty cycle, voltage, and servo-related variables in EDM and Wire EDM processes.

Several studies have demonstrated that discharge current and pulse-on time are among the dominant factors affecting EDM performance. Puertas et al. [7] investigated the influence of current intensity, pulse-on time, and pulse-off time on surface quality and dimensional precision, concluding that current intensity has a major effect on surface roughness. El-Taweel et al. [8] studied EDM using Al-Cu-Si-TiC composite electrodes and found peak current to be the most influential parameter for both material removal rate and tool wear rate. Similarly, Majumder et al. [9] reported that supply current has the strongest effect on material removal rate, while pulse-on time significantly influences electrode wear.

Taguchi-based and response surface methodology approaches have also been widely applied for EDM parameter optimization. Lajis et al. [10] used the Taguchi method for EDM of tungsten carbide and showed that peak current mainly affects electrode wear and surface roughness, while pulse-on time strongly influences material removal rate. Sohani et al. [11] applied response surface methodology to study tool shape, discharge current, pulse-on time, pulse-off time, and tool area, demonstrating that interactions between discharge current and pulse-on time significantly affect both material removal rate and tool wear rate.

Other works extended EDM optimization toward multi-objective and hybrid techniques. Datta et al. [12] combined response surface methodology with Grey-Taguchi analysis for Wire EDM and optimized material removal rate, surface roughness, and kerf width simultaneously. Subrahmanyam et al. [13] applied Grey-Taguchi optimization to Wire EDM using multiple parameters, including pulse-on time, pulse-off time, spark voltage, wire tension, servo feed, and flushing pressure. These studies show that EDM optimization is often a multi-response problem, where improving productivity may negatively affect surface integrity or tool wear.

Research has also shown that parameter importance depends on the selected material, electrode, and target response. Patel et al. [14] found pulse-on time to be the dominant factor for surface roughness in EDM of ceramic composites, while Chattopadhyay et al. [15] showed that peak current, pulse-on time, and electrode rotation affect material removal rate, surface roughness, and electrode wear differently. Nikalje et al. [16] further demonstrated that the optimal parameter levels for material removal rate and relative wear ratio may differ from those for surface roughness and tool wear.

Overall, previous studies confirm that EDM performance is governed by a complex and highly interacting parameter space. Most approaches focus on direct optimization using statistical or evolutionary methods, while less attention is given to identifying compact, task-specific parameter subsets. Therefore, this study applies feature selection methods to determine the most relevant parameters for regime separation and efficient process control.

III. DATASET AND EXPERIMENTAL SETUP

A. Input Parameters

The dataset analysed in this study is derived from controlled industrial EDM experiments carried out in the context of manufacturing. The primary goal of the experimental design was to investigate the effect of key technological parameters on machining performance while reducing the complexity of the parameter space to enable more efficient experimentation and robust process optimization.

The input space consists of six essential EDM technological parameters and are anonymized and referred to collectively as Technological Parameters X. This abstraction ensures that proprietary process knowledge remains protected, while still allowing for a generalizable and methodologically valid analysis. From a modelling perspective, the problem can therefore be interpreted as operating within a six-dimensional parameter space, where the identification of the most influential parameters is a central objective.

B. Process Description and Measured Variables

The EDM process is structured into three consecutive machining stages, namely setup, processing, and completion. For each of these stages, detailed operational data were recorded, enabling a phase-specific analysis of machining behaviour. The collected outputs include both time-related and tool wear-related measurements for each stage, as well as their cumulative values over the entire machining process.

Despite the availability of tool wear information, the present study focuses exclusively on machining time as the primary performance indicator. This decision is motivated by the relatively low cost of the tool used in the process, which reduces the practical relevance of wear in the given industrial setting. Consequently, the analysis is focus on time-based metrics, including individual stage durations and total machining time, which are considered more critical for process efficiency and optimization.

Each experiment contains over hundred machining records, where each record corresponds to a complete EDM operation defined by a specific combination of technological parameters. As this represents a relatively small dataset for machine learning, only methods that can reliably operate under limited data availability are considered. Therefore, two such methods are proposed and evaluated in Section V.

IV. PROBLEM FORMULATION

The machining experiments are organized according to machining time, which provides a natural ordering reflecting process efficiency and enables the formulation of subsequent learning tasks A and B.

A. Task A) Classification of Machining Regimes

The first level of analysis is defined as a binary classification problem, where each machining instance is assigned to one of two categories: high-speed machining or low-speed machining. This classification is based on the ordered time values and was determined in collaboration with domain experts to reflect practically meaningful performance thresholds.

The main objective of this task is to identify which technological parameters are most influential in separating these two machining regimes. By analysing the contribution of individual features to class separation, it becomes possible to reduce the dimensionality of the input space and to optimize the parameters that are most critical for achieving high machining efficiency.

B. Task B) High-Speed Regime Refinement

The second level of analysis focuses exclusively on the subset of machining instances classified as high-speed. In this reduced dataset, the objective shifts from general classification to a more refined exploration of parameter influence within the high-performance regime. Specifically, the goal is to identify the most favourable parameter configurations among already efficient machining conditions, rather than distinguishing between efficient and inefficient cases.

This refined problem formulation allows for a more detailed investigation of parameter importance in a constrained region of the parameter space. The definition of the high-speed subset, as well as the criteria for identifying optimal configurations within it, were also established in collaboration with domain experts to ensure practical relevance. By excluding low-speed records, the analysis becomes more sensitive to subtle differences in performance, enabling the identification of parameter combinations that contribute to further improvements in machining efficiency.

V. METHODOLOGY

Given the relatively limited number of machining records available in each experiment, two feature selection methods suitable for small-data scenarios are employed: Adaptive Hybrid Feature Selection (AHFS) [17–18], which performs task-driven feature selection, and Model Input–Output Configuration Search with Embedded Feature Selection (MICS-EFS) [20–21], which provides a task-independent assessment of feature importance.

A. Adaptive Hybrid Feature Selection (AHFS)

The first applied method is the Adaptive Hybrid Feature Selection (AHFS) algorithm [17–18], which is designed to identify the most informative subset of input features in a task-driven manner. AHFS follows an iterative forward selection strategy, starting from an empty feature set and incrementally extending it by adding one feature at a time. At each iteration, candidate features are identified using a combination of state-of-the-art information-theoretic and correlation-based measures, ensuring that multiple perspectives on feature relevance are considered.

The final selection among candidate features is guided by a predictive model, specifically an artificial neural network trained using the Levenberg–Marquardt optimization algorithm [19]. As a result, the AHFS method produces an ordered list of features ranked according to their contribution to the prediction task.

B. Model Input-Output Configuration Search with Embedded Feature Selection (MICS-EFS)

The second applied method is the Model Input-Output Configuration Search with Embedded Feature Selection (MICS-EFS) approach [20–21], which provides a task-independent assessment of feature importance. Unlike AHFS, this method does not rely on a predefined target variable associated with specific machining stages but instead aims to identify features that best represent the overall structure of the input space.

MICS-EFS is also based on a sequential forward search but combined with a reconstruction-based evaluation mechanism. In each iteration, the algorithm selects the feature that most effectively contributes to representing the remaining features. The evaluation of this representation is performed using an encoder–decoder model. The encoder maps the selected features into a latent representation, while the decoder attempts to reconstruct the full feature set from this reduced representation. Through this iterative process, the method produces a general feature ranking that reflects internal relationships among the input variables rather than task-specific predictive performance.

VI. RESULTS AND DISCUSSIONS

A. Results for Task A: Classification of Machining Regimes

The feature rankings obtained by the AHFS and MICS-EFS feature selection methods, are presented in Table 1.

Table 1. Feature rankings for classification of machining regimes

Method	Time	Technological parameters					
		1	2	3	4	5	6
AHFS	Total	1	2	3	4	5	6
	Setup	1	4	3	5	6	2
	Proc.	1	6	3	4	5	2
	Comp.	1	2	3	5	6	4
MICS	–	1	2	3	5	4	6

The AHFS results show that the first technological parameter is consistently the most important across all machining stages, indicating its dominant role in separating machining regimes. In contrast, the ranking of the remaining parameters varies between the setup, processing, and completion phases, highlighting the stage-dependent nature of parameter importance. The MICS-EFS ranking, which is independent of the machining stages, shows the strongest similarity to the total and completion time rankings, suggesting that the global structure of the parameter space is mainly influenced by later process stages.

The predictive performance of the AHFS models is shown in Fig. 2 and Fig. 3 in terms of accuracy and error.

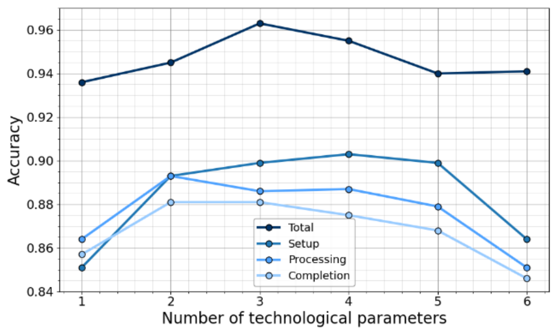

Fig. 2. Accuracy of classification of machining regimes

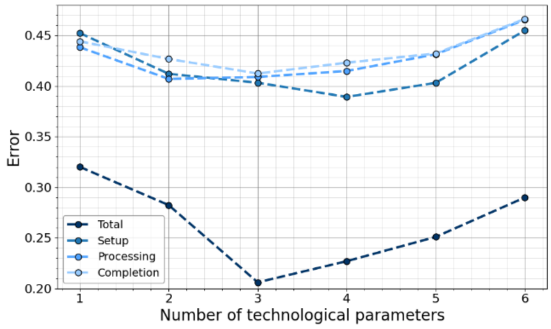

Fig. 3. Error of classification of machining regimes

The proposed approach achieved a classification accuracy of up to 96% in separating high-speed and low-speed machining regimes. Both figures indicate the presence of an optimal number of technological parameters, beyond which model performance deteriorates. The optimal technological parameter subset sizes for AHFS are found to be 3, 4, 2, and 2 for total, setup, processing, and completion time, respectively. The MICS-EFS method also identifies three technological parameters as optimal, confirming that a reduced subset of parameters is sufficient for effective classification.

B. Results for Task B: High-Speed Regime Analysis

The feature rankings obtained for the high-speed regime are summarized in Table 2 for both AHFS and MICS-EFS methods across the different machining stages.

Table 2. Feature rankings for high-speed regime analysis

Method	Stage	Technological parameters					
		1	2	3	4	5	6
AHFS	Total	1	4	6	3	5	2
	Setup	1	5	6	3	2	4
	Proc.	1	3	5	6	4	2
	Comp.	1	4	5	2	6	3
MICS	–	1	4	6	3	5	2

Similarly to Task A, the results indicate that the Technological Parameter 1 plays a dominant role across all stages, confirming its overall importance even within the high-speed domain. However, the ranking of the remaining parameters shows variation between the setup, processing, and completion stages, further supporting the stage-dependent nature of parameter importance. The MICS-EFS ranking, which is independent of the machining stages, shows full agreement with the total time ranking, indicating that the global structure of the parameter space in the high-speed regime is primarily reflected by the total machining performance.

The predictive performance of the AHFS models is presented in Fig. 4 and Fig. 5 in terms of classification accuracy and error.

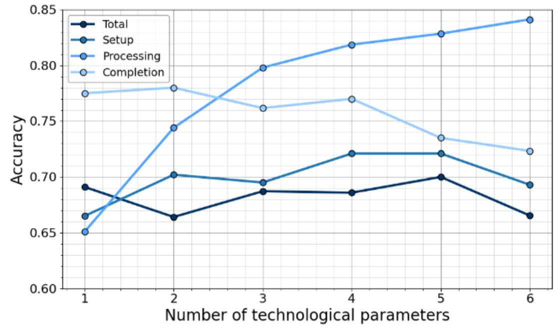

Fig. 4. Accuracy of classification of high-speed regimes

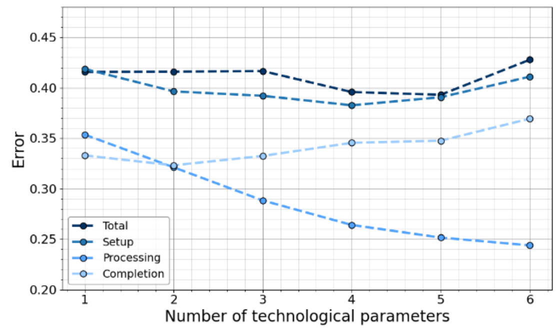

Fig. 5. Error of classification of high-speed regimes

Compared to Task A, the classification of high-speed regimes represents a more challenging problem, as reflected by the lower accuracy values, ranging approximately between 75% and 85%. This is expected, since the task focuses on distinguishing among already efficient machining conditions, where differences are inherently smaller and more difficult to capture.

The results also indicate the presence of an optimal number of technological parameters for most machining stages. The optimal technological parameter subset sizes for AHFS are found to be 5, 4, and 2 for total, setup, and completion stages, respectively, beyond which model performance deteriorates. In contrast, for the processing stage, the best performance is achieved when all six parameters are included, however, their ranked importance still provides valuable insight from a process control perspective, enabling prioritization of parameter tuning and more structured control strategies. A similar observation holds for the MICS-EFS method, where the optimal performance is also obtained using all available technological parameters.

These findings suggest that, in the high-speed regime, certain stages require a more complete representation of the parameter space to capture subtle performance differences, while others can still be effectively described using a reduced subset of technological parameters. Overall, the results highlight the increased complexity of fine-grained process analysis and further emphasize the importance of stage-dependent feature selection in EDM process optimization.

C. Comparison of Task A and B

The results of Task A (Classification of Machining Regimes) and Task B (High-Speed Regime Analysis) provide complementary insights into the role of technological parameters in EDM process analysis. While both tasks aim to identify influential parameters, they differ in their level of complexity and the nature of the classification problem.

A key similarity between the two tasks is that Technological Parameter 1 consistently emerges as the most important technological parameter across all machining stages and both feature selection methods. This confirms its dominant and stable influence on machining performance, regardless of whether global regime separation or fine-grained differentiation is considered.

However, notable differences can be observed in the importance of the remaining parameters. In Task A, the classification between high-speed and low-speed regimes results in relatively stable and interpretable rankings, with clear optimal technological parameter subset sizes. In contrast, Task B exhibits higher variability in parameter rankings across stages, indicating that parameter importance becomes more sensitive and stage-dependent when analysing only high-speed regimes. This increased variability reflects the more subtle and complex nature of differences within already optimized machining conditions.

The comparison of model performance further highlights the distinction between the two tasks. Task A achieves a high classification accuracy of up to 96%, demonstrating that global regime separation is relatively well-defined. In contrast, Task B yields lower accuracy values in the range of 75–85%, confirming that distinguishing between high-speed machining instances is inherently more challenging due to smaller performance differences.

Differences are also evident in the optimal number of technological parameters. While Task A consistently shows clear optimal technological parameter subset sizes across all stages, Task B reveals that a reduced technological parameter set is not always sufficient. In particular, for the processing stage and the MICS-EFS method, the best performance is achieved using all six parameters, suggesting that a more complete representation of the parameter space is required to capture subtle variations in high-speed machining behaviour.

Overall, the comparison demonstrates that global classification tasks benefit from dimensionality reduction and highlight dominant parameters, whereas fine-grained analysis within high-performance regimes requires a more detailed and comprehensive parameter representation. These findings support the use of a two-stage analytical framework, where coarse classification is followed by refined analysis, enabling both efficient parameter reduction and detailed process understanding.

VII. CONCLUSIONS AND OUTLOOK

This study presented a data-driven framework for analysing EDM process parameters using feature selection methods applied to an industrial dataset. The results demonstrated that the original six-dimensional parameter space can be effectively reduced while maintaining high classification performance. In Task A, a classification accuracy of up to 96% was achieved, with clearly identifiable optimal technological parameter subsets. In contrast, Task B highlighted the increased complexity of fine-grained analysis within the high-speed regime, where lower accuracy (75–85%) and higher sensitivity to parameter selection were observed.

The findings revealed that while one technological parameter consistently dominates across all conditions, the importance of the remaining parameters is strongly stage dependent. Furthermore, the results showed that reduced technological parameter sets are sufficient for global classification tasks, whereas more detailed parameter representations are required for refined analysis. In addition, the reduced parameter space enables simpler process control with fewer variables to be optimally tuned and significantly decreases the number of required experiments in design of experiments (DoE) for process optimisation. Future work will focus on the detailed evaluation of both Task A and Task B for tool wear analysis.

VIII. ACKNOWLEDGMENTS

The research was supported by the European Commission (EC) through AI EDIH Hungary (Project ID: 101120929). On behalf of Project "Testing Machine Learning Algorithms", we thank for the usage of HUN-REN Cloud (https://science-cloud.hu/) that significantly helped us achieving the results published in this paper.

REFERENCES

[1] **Ho, K. H.; Newman, S. T.** State of the art electrical discharge machining (EDM). *International Journal of Machine Tools and Manufacture*, 2003, pp. 1287-1300.

[2] **Abbas, N. M., Solomon, D. G., & Bahari, M. F.** A review on current research trends in electrical discharge machining (EDM). *International Journal of machine tools and Manufacture*, 2007, pp. 1214-1228.

[3] **Antar, M., Chantzis, D., Marimuthu, S., & Hayward, P.** High speed EDM and laser drilling of aerospace alloys. *Procedia CIRP*, 2016, pp. 526-531.

[4] **Kliuev, M., Boccadoro, M., Perez, R., Dal Bó, W., Stirnimann, J., Kuster, F., & Wegener, K.** EDM drilling and shaping of cooling holes in Inconel 718 turbine blades. *Procedia CIRP*, 2016, pp. 322-327.

[5] **Kliuev, M., & Wegener, K**. Method of machining diffusors in Inconel 718 turbine blades for film cooling using EDM drilling and shaping. *Procedia CIRP*, 2020, pp. 511-515.

[6] **Singh, V., & Pradhan, S. K.** Optimization of EDM process parameters: a review. *International Journal of Emerging Technology and Advanced Engineering*, 2014, pp. 345-355.

[7] **Puertas, I., & Luis, C. J.** A study on the machining parameters optimisation of electrical discharge machining. *Journal of Materials Processing Technology*, 2003, pp. 521-526.

[8] **El-Taweel, T. A**. Multi-response optimization of EDM with Al–Cu–Si–TiC P/M composite electrode. *The International Journal of Advanced Manufacturing Technology*, 2009, pp. 100-113.

[9] **Majumder, A.** Study of the effect of machining parameters on material removal rate and electrode wear during Electric Discharge Machining of mild steel. *Journal of Engineering Science and Technology Review*, 2012, pp. 14-18.

[10] **Lajis, M. A., Radzi, H. C. D. M., & Amin, A. K. M. N**. The implementation of Taguchi method on EDM process of tungsten carbide. *European Journal of Scientific Research*, 2009, pp. 609-617.

[11] **Sohani, M. S., Gaitonde, V. N., Siddeswarappa, B., & Deshpande, A. S**. Investigations into the effect of tool shapes with size factor consideration in sink electrical discharge machining (EDM) process. *The International Journal of Advanced Manufacturing Technology*, 2009.

[12] **Datta, S., & Mahapatra, S. S.** Modeling, simulation and parametric optimization of wire EDM process using response surface methodology coupled with grey-Taguchi technique. *International journal of engineering, science and Technology*, 2010, pp. 162-183.

[13] **Subrahmanyam, S. V., & Sarcar, M. M. M.** Evaluation of optimal parameters for machining with wire cut EDM using Grey-Taguchi method. *International Journal of Scientific and Research Publications*, 2013, pp. 1-9.

[14] **Patel, K. M., Pandey, P. M., & Rao, P. V.** Determination of an optimum parametric combination using a surface roughness prediction model for EDM of Al2O3/SiCw/TiC ceramic composite. Materials and Manufacturing Processes, 2009, pp. 675-682.

[15] **Chattopadhyay, K. D., Verma, S., Satsangi, P. S., & Sharma, P. C.** Development of empirical model for different process parameters during rotary electrical discharge machining of copper–steel (EN-8) system. *Journal of Materials Processing Technology*, 2009, pp. 1454-1465.

[16] **Nikalje, A. M., Kumar, A., & Srinadh, K. S.** Influence of parameters and optimization of EDM performance measures on MDN 300 steel using Taguchi method. The *International Journal of Advanced Manufacturing Technology*, 2013, pp. 41-49.

[17] **Viharos, Z. J., Kis, K. B., Fodor, Á., & Büki, M. I**. Adaptive, hybrid feature selection (AHFS). *Pattern Recognition*, 2021, 116, 107932.

[18] **Gercuj, B., Hoang, A. T., & Viharos, Z. J.** Towards a Modular Redesign of the Scalable Adaptive Hybrid Feature Selection Algorithm. *In 2025 IEEE 23rd Jubilee International Symposium on Intelligent Systems and Informatics (SISY)*, 2025, pp. 369-374

[19] **Moré, J. J.** The Levenberg-Marquardt algorithm: implementation and theory. *In Numerical analysis: Proceedings of the Biennial Conference held at Dundee*, 2006, pp. 105-116

[20] **Hoang, A. T., & Viharos, Z. J**. Model input-output configuration search with embedded feature selection for sensor time-series and image classification. *IEEE Access*, 2025.

[21] **Hoang, A. T., & Viharos, Z. J.** Robust superiority of the MICS-EFS configuration search algorithm through modular extensions of complex neural architectures," *In 2025 IEEE 19th International Symposium on Applied Computational Intelligence and Informatics (SACI),* 2025