Data Science Tool “MELSOFT MaiLab”

Authors: *Shota Nakano** and *Hiroshi Otsu**

In recent years, the importance of data utilization in corporate activities has been increasing, and greater efforts are being made to improve production sites based on data in the manufacturing industry. However, advanced data analysis beyond mere data collection and visualization has yet to spread due to the lack of personnel with specialized knowledge, such as data scientists.

Against this background, Mitsubishi Electric has developed the data science tool “MELSOFT MaiLab” as a solution that enables advanced data analysis without requiring specialized knowledge and seamlessly applies the analysis results to production sites.

MELSOFT MaiLab incorporates AI technology such as deep learning, which has progressed remarkably in recent years, and has a function that enables automatic learning by AI with simple operations without requiring expertise. This allows MELSOFT MaiLab itself to act like a dedicated AI data scientist, providing strong support for on-site problem solving and improvement.

It also has a function for customizing the diagnostic models automatically learned by AI, thereby allowing a flexible response when each production site requires a unique response or when a data analysis expert wants to improve the models.

MELSOFT MaiLab facilitates data utilization in the manufacturing industry and enables anyone to improve production sites by using data. This will contribute to the generalization of data analysis and the improvement of corporate competitiveness.

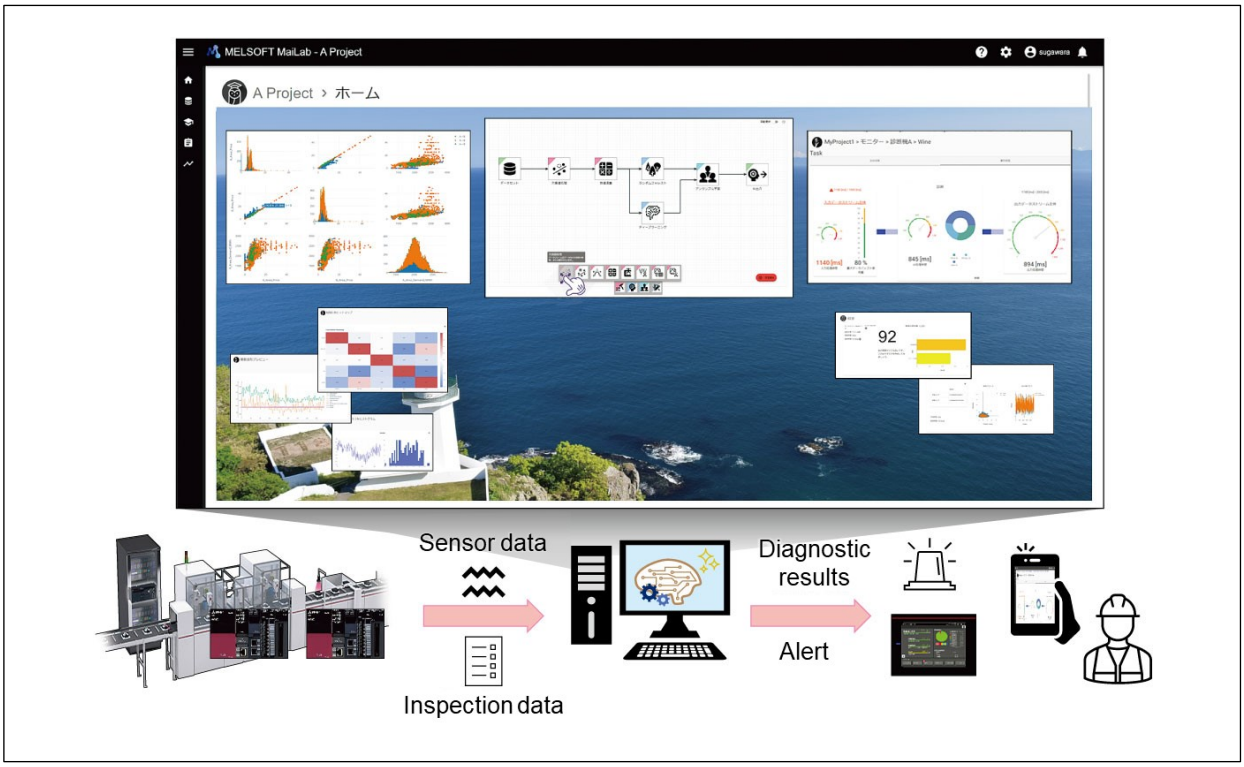


Image of Using MELSOFT MaiLab

MELSOFT MaiLab first collects and accumulates sensor data and inspection data from the production site, which AI automatically analyzes, and creates a diagnostic model. When production site data is input to the model during operation, AI performs diagnosis and feeds back the diagnostic results and alerts. The software automatically performs data visualization, diagnostic model creation, and diagnostic model evaluation, etc., allowing users to perform these data tasks by simple operations through an intuitive UI (User Interface).

* *Nagoya Works*

1. Introduction

Many manufacturing companies are trying to utilize data to improve their production sites. However, about half of companies are interested but have not yet started to use data, due to a lack of data analysis personnel and skills, and so are seeking tools that can be easily introduced and utilized. In addition, while they want tools that can be used without specialized knowledge, they also want customizability, such as for linking with their own programs to solve their own problems.

The data science tool MELSOFT MaiLab acts like a data scientist, such as creating diagnostic models and proposing production improvements, even if the person in charge of improvement does not have expertise in data analysis. It also has an intuitive UI and customizable functions using proprietary programs to encourage the use of data at production sites.

2. Overview of MELSOFT MaiLab

The general flow for improving the production site by utilizing data consists of: “collecting data” from sequencers and sensors; “analyzing the collected data to create a diagnostic model”; and “applying the created model to the site for operation” (Fig. 1). The phase of “analyzing the collected data to create a diagnostic model” has five more steps, each of which takes time and effort and requires knowledge of data analysis such as statistics, making it difficult to use data.

MELSOFT MaiLab uses AI to automate data analysis and diagnostic model creation, enabling production sites to be improved through data utilization without specialized knowledge. In addition, MELSOFT MaiLab is the software that supports consistently not only data analysis and diagnostic model creation, but also data collection for analysis, and on-site application and operation monitoring of the diagnostic model.

3. Characteristics of MELSOFT MaiLab

3.1 Simple analysis

One of the characteristics of AI-based analysis support is the AI automatic learning function using AutoML (Automated Machine Learning), which enables anyone to easily analyze data and create a diagnostic model.

This software automatically creates the optimal diagnostic model for achieving the purpose simply by answering questions about the purpose of using data such as “detecting anomalies” and “predicting the future” and information such as “availability of training data”(Fig. 2).

In order to measure the quality of the created diagnostic model, the software performs scoring based on an index and plots the evaluation results for the test data. It not only generates the diagnostic model, but also shows learning results and accuracy to support decision-making on using the model (Fig. 3).

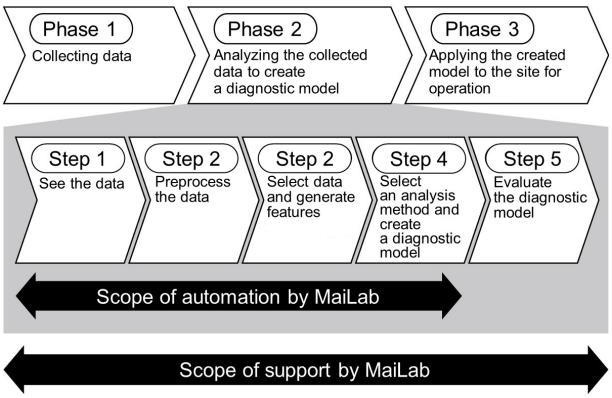


Fig. 1 Flow of improvement through data utilization

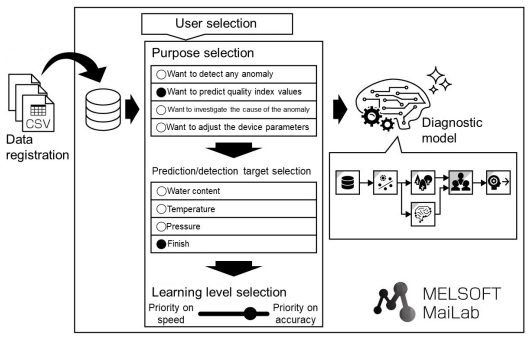


Fig. 2 Image of AI automatic learning procedure

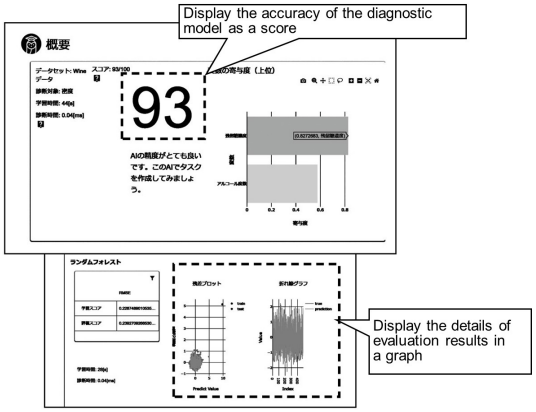


Fig. 3 Learning result display screen

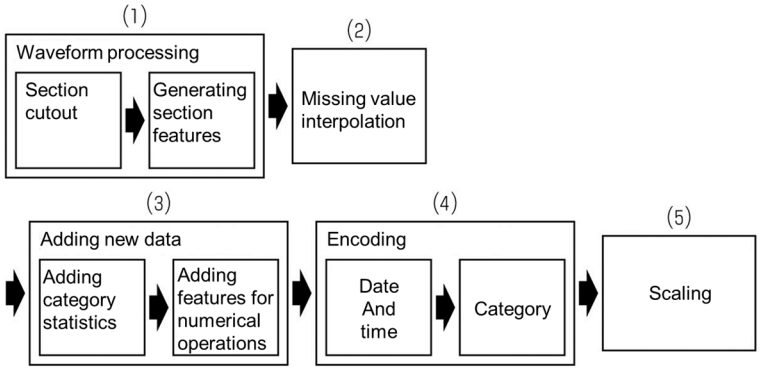


Fig. 4 Flow of data preprocessing

3.1.1 Advanced data preprocessing functions

To improve the accuracy of the diagnostic model, data preprocessing is important, such as cleaning up the target data, narrowing it down to the data required for analysis, and identifying the characteristics of the data. MELSOFT MaiLab automatically performs the following five preprocessing steps to optimize the data (Fig. 4).

(1) Waveform processing
For waveform data that changes over time such as sensor measurement data, the data is segmented into sections and cut out, and the cut-out data is converted into features.

(2) Missing value interpolation
Appropriate interpolation processing is performed for missing values (such as blanks when sensing is not possible).

(3) Adding new data
Variables that have a large impact on the diagnosis are automatically identified, and are then converted to statistics and added as new features. In addition, variables with high importance are calculated and added as new features.

(4) Encoding
Data that cannot be learned in its original data format is converted into a format that can be learned while maintaining the characteristics of the data.

(5) Scaling
To ensure proper learning and to facilitate evaluation of learning results, the data are converted to an equivalent scale.

3.1.2 Effective and accurate learning

Automatic learning is performed using preprocessed data. Instead of just automatic processing, tuning is performed for effective learning, and multiple methods are combined to enable highly accurate prediction.

MELSOFT MaiLab performs the following three processes to automate the entire learning process and generate better diagnostic models.

(1) Optimizing hyperparameters
Hyperparameters are set values that control the behavior of the analysis method. Since prediction results and accuracy vary depending on the settings, MELSOFT MaiLab performs search and optimization for each analysis method.

(2) Equipped with various analysis methods
In addition to conventional statistical methods, machine learning techniques such as deep learning are provided, allowing for highly expressive learning even from highly complex data.

(3) Ensemble learning
Ensemble learning is a method that combines the prediction results of multiple analysis methods to improve prediction accuracy. Since the calculation method differs depending on the analysis method, even if the same data is used for learning or prediction, the results will differ for each analysis method. Predictions based on a single method may result in biased results depending on the characteristics of the data and method. Through ensemble learning, the tendencies of strengths and weaknesses of each method are learned, which enables more accurate predictions than those made with a single method (Fig. 5).

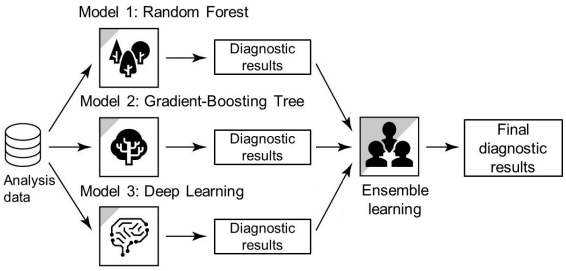


Fig. 5 Image of ensemble learning

3.2 Free customization of diagnostic models

The automatically created diagnostic models can be checked and edited with a dedicated editor, from data input to prediction result output (Fig. 6). The icons in the figure represent functional blocks that package data processing methods and analysis methods, and the models can be easily edited simply by arranging and connecting the blocks by using a mouse.

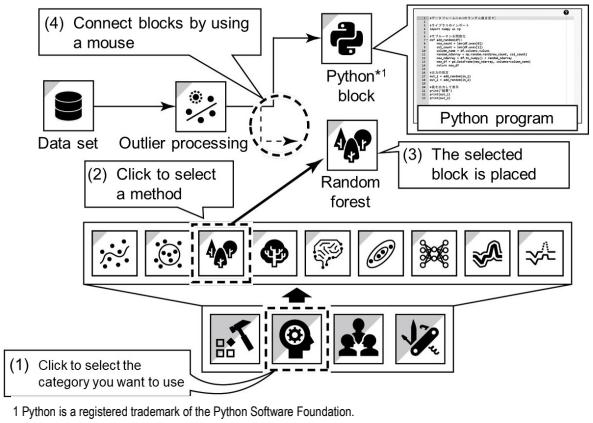
The automatically generated diagnostic models can be improved by the user performing their own preprocessing such as data processing. Instead of relying on the software, users can also create diagnostic models by using their own data analysis knowledge, such as creating the preprocessing and diagnostic methods by themselves from scratch.

In addition, there are functional blocks that allow users to incorporate their own Python code, which makes it possible to improve the accuracy of diagnostic models by linking them with the users' original processing (Fig. 7).

3.3 Convenient data utilization

In order to improve production sites by utilizing data, it is necessary to apply diagnostic models to the sites and realize condition monitoring and even predictive detection based on data generated in real time at the site. Common data analysis software rarely has a function for operating a diagnostic model; usually, a separate system must be built. MELSOFT MaiLab software enables both creation and operation of diagnostic models in a single software package, which can be deployed on the same PC or another PC and operated as is (Fig. 8).

It also has a data linkage function with basic software of Mitsubishi Electric's sequencer MELSEC and Edgecross, which eliminates the need to construct a separate data acquisition system or diagnostic system, thereby enabling faster start-up time and reducing costs.



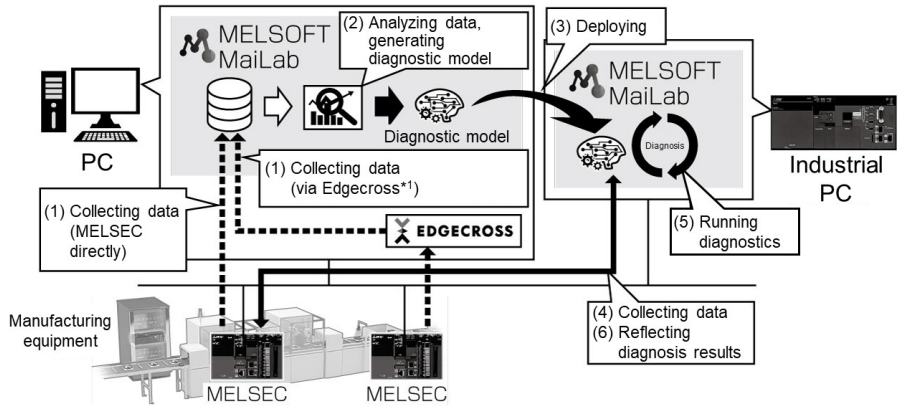
1 Python is a registered trademark of the Python Software Foundation.

Fig. 6 Model editing with the editor

Functional blocks for preprocessing					
	Missing value processing		Outlier processing		Scaling
	Numerical operation		Dimensional compression		Section generation (Waveform cutout)
	Section feature generation		Subsection shaping		
Functional blocks for analysis methods					
	Deep learning		Gradient-boosting decision tree		Random forest
	Auto encoder		K-nearest neighbor method		Similar waveform recognition
	Multiple regression		MT method		Guard band

MT : Mahalanobis−Taguchi

Fig. 7 Functional block list for analysis



1 Edgecross is a registered trademark of the Edgecross Consortium.

Fig. 8 Example of system configuration

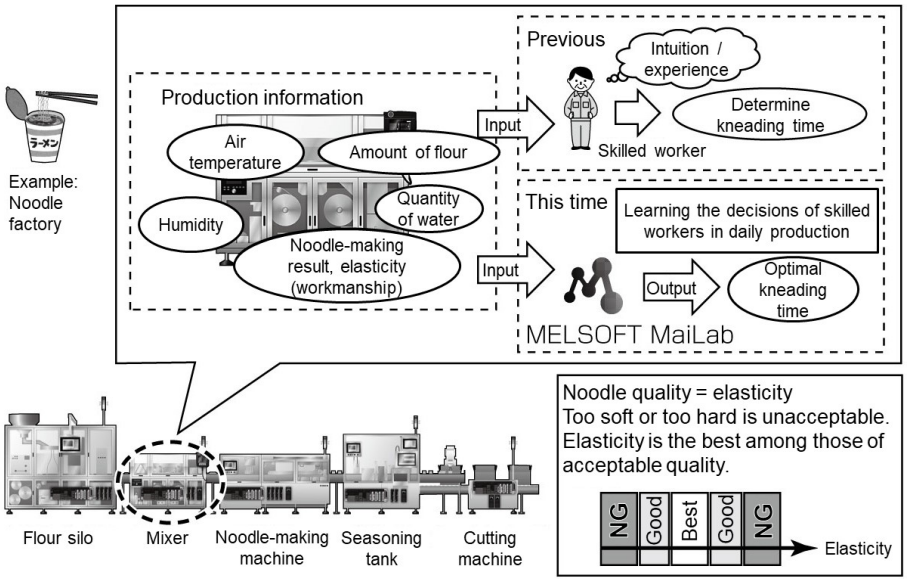


Fig. 9 Use case of optimal control parameter proposal

4. Example of Utilization

MELSOFT MaiLab allows non-data scientists to create diagnostic models using technologies such as deep learning and apply them to actual sites. As an example of data utilization at production sites, this section describes how to apply diagnostic models to inherit and pass on skilled workers' know-how, which is an issue faced by many manufacturing companies.

4.1 Proposal of optimal control parameters

Skilled workers can determine and set the desired control parameters at that time based on their own intuition and experience. However, this know-how is tacit, and is lost when skilled workers retire.

MELSOFT MaiLab can create a diagnostic model that proposes control parameter settings similar to those chosen by skilled workers by learning both manufacturing information and control parameters determined by skilled workers as a set (Fig. 9). This enables even inexperienced workers to make settings similar to those of skilled workers.

4.2 Estimation of the causes of anomalies

When an equipment failure or product defect occurs, the cause is estimated by investigating past events after the anomaly. However, workers with little experience of dealing with anomalies may have little know-how and take a long time to investigate.

MELSOFT MaiLab can estimate the cause of an anomaly based on the circumstances under which it occurred, by using AI to learn information about what types of anomalies occurred in the past and under what circumstances. This helps to investigate the cause, which usually requires experience and know-how, and to shorten the investigation time. As a result, equipment downtime can be reduced and productivity improved.

5. Conclusion

This paper described the development background and characteristics of MELSOFT MaiLab, which provides total support for data utilization at production sites, and the technologies its uses.

Data analysis and AI technologies are constantly advancing, and as the number of powerful technologies continues to increase, the shortage of data analysis personnel and skills is becoming more serious. Mitsubishi Electric will continue to contribute to solving issues and improving productivity at manufacturing sites by enhancing new technologies and expanding the scope of their application.