Data Science Tool "MELSOFT MaiLab"

Author: *Shota Nakano**

1. Introduction

Many manufacturing companies would like to leverage data to improve production sites. Although a majority of these companies have shown interest in data-driven enhancements, few have been able to shift this interest into action. A major obstacle inhibiting these businesses from taking action is a lack of knowledge and human resources skilled in data analysis within their organizations. In other words, these companies need a tool to easily adopt and use data analysis. While the hope is for a tool that can be used without any specialized knowledge, companies still want customizability, including compatibility with their own programs used to overcome unique problems.

Mitsubishi Electric developed the Data Science Tool MELSOFT MaiLab as a solution for a tool that can act as the data scientist, generate diagnosis models, and propose production enhancements, which eliminates the data science knowledge barrier for employees in charge of production improvements. The intuitive User Interface (UI) and compatibility with independent programs facilitate customization and encourage data-driven production sites.

2. MELSOFT MaiLab Overview

The overall process to enhance production sites by leveraging data consists of collecting data from programmable controllers and sensors, analyzing the data that has been collected, creating a diagnosis model, and embedding that diagnosis model into the production process (Fig. 1). The phase to analyze the data and create a diagnosis model breaks down into five additional steps. Each of these steps is tedious and takes time while also requiring knowledge about statistics or other data analysis skills. These necessities make the barrier to start using data quite high.

MELSOFT MaiLab takes advantage of AI to automate the data analysis and diagnosis model creation process, effectively eliminating the knowledge barrier to realize data-driven enhancements to production sites. The software goes beyond only data analysis and diagnosis model creation as well to support the entire process from data collection and analysis to use and operational monitoring of the diagnosis model in production.

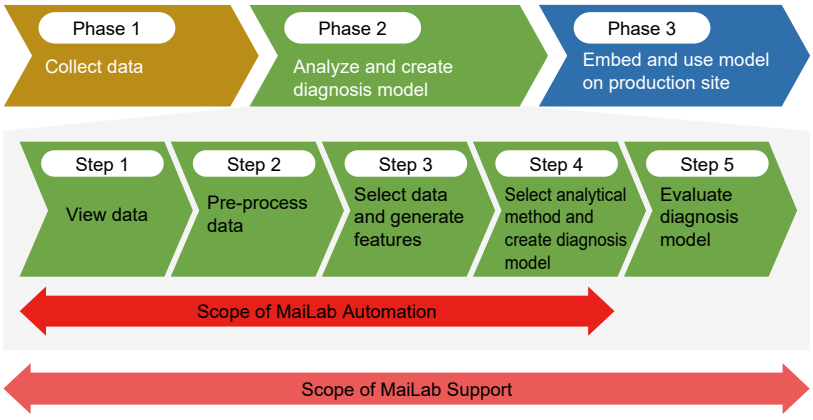


Fig. 1 Flow of data-driven enhancements

**Nagoya Works*

3. MELSOFT MaiLab Features

3.1 Simple analysis

One AI analysis support feature is the Automated Machine Learning (AutoML) function. AutoML achieves simple data analysis and diagnosis model creation that anyone can use.

Whether the purpose of the model is detecting Abnormalities, predicting the future, or inputting training or other data, MELSOFT MaiLab can automatically generate an ideal diagnosis model to accomplish those objectives by simply asking the users questions (Fig. 2).

The software also provides a score based on the parameters and graphs the evaluation results for the test data to assess the quality of the diagnosis model generated by the AI system. This score not only supports the creation of diagnosis models but also summarizes the training results and accuracy in addition to helping users determine whether to embed the model into the production process (Fig. 3).

3.1.1 State-of-the-art data pre-proccessing functions

Data pre-processing before creating a diagnosis model is vital to improve the accuracy. This includes properly organizing the relevant data, narrowing the data set necessary for an analysis, and assessing the data quality. MELSOFT MaiLab offers five functions to automatically process and optimize data (Fig. 4):

(1) Waveform data processing

This function extracts interval waveform data as the sensor measurements change with time and performs featurization of the data that has been extracted.

(2) Imputation of missing data

This imputation process properly inserts missing data (values left blank due to a lack of sensor data).

(3) Addition of new data

This function automatically identifies variables that largely impact diagnostic processes to create and input these variables as new features. The software also calculates the higher priority between those variables to add as new features.

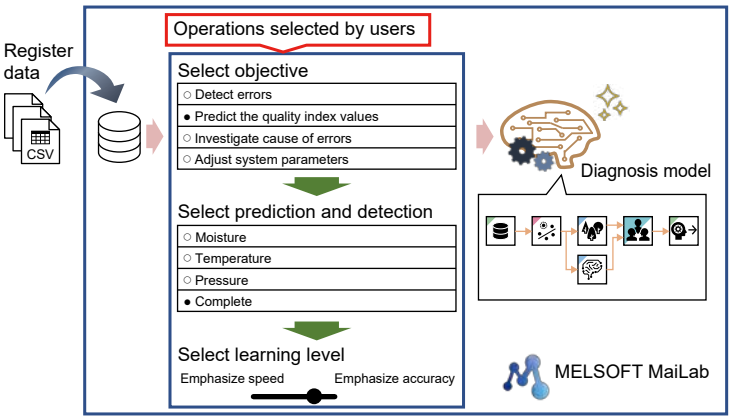


Fig. 2 Steps of automated AI training

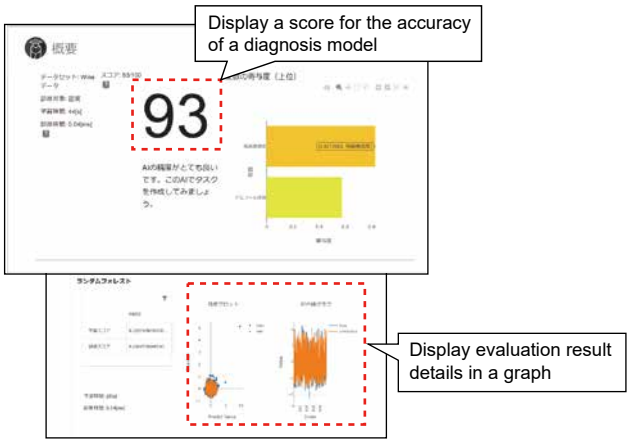


Fig. 3 Screen displaying learning results

(4) Encoding

The encoding process converts data in a format that cannot be used for training into a data format that can be used for training while retaining the features of the original data.

(5) Scaling

The scaling process converts data into an equivalent scale to properly use for training to make the evaluation of training results easier.

3.1.2 Effective and highly accurate machine learning

The software automatically executes machine learning by analyzing data in advance. The data pre-processing does not automatically process the data but rather tunes the data for more effective machine learning and combines multiple algorithms to execute machine learning using highly accurate predictions.

MELSOFT MaiLab provides three functions to mine data and generate higher quality diagnosis models by automating the entire machine learning process.

(1) Hyperparameter optimization

A hyperparameter is a parameter used to control the behavior of each learning algorithm. MELSOFT MaiLab explores and optimizes each learning algorithm because the prediction results and accuracy vary according to the parameter settings.

(2) Various learning algorithms

MELSOFT MaiLab offers deep learning and a variety of other machine learning processes in addition to the typical statistical process. These new machine learning processes can perform training with higher quality representations through highly complex data.

(3) Ensemble learning

The ensemble learning process combines prediction results from multiple learning algorithms to improve predictive performance. Each learning algorithm obtains different results even when learning and making predictions on the same data sets because the calculation methods differ for each algorithm. A single learning algorithm presents bias in the results due to the data and algorithm features when making predictions. Ensemble learning learns the pros and cons of each learning algorithm to realize more highly accurate predictions than any single algorithm alone (Fig. 5).

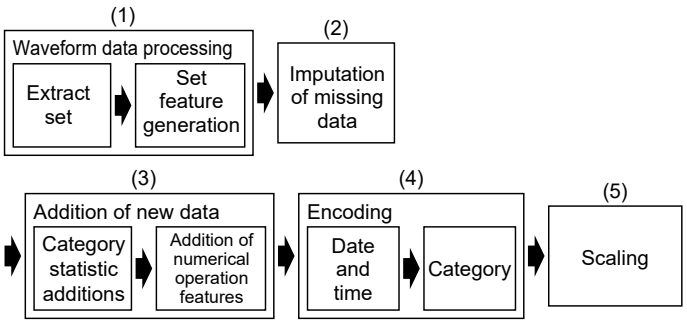


Fig. 4 Flow of data pre-processing

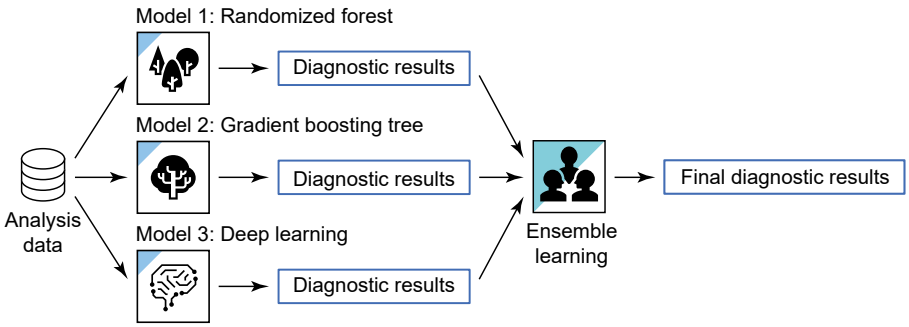


Fig. 5 Illustration of ensemble learning

3.2 Flexible diagnosis model customization

Users can use a dedicated editor to verify and edit diagnosis models automatically generated by the system from data input through predictive results output (Fig. 6). The icons in the figure indicate the function blocks packaged for data processing and analysis methods. Users can arrange and connect these blocks using mouse operations to easily edit a diagnosis model.

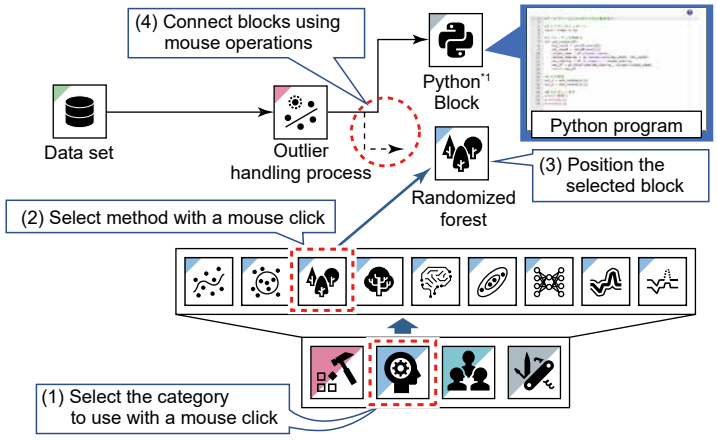
The user can also enhance data and other pre-processing to improve the diagnosis model automatically generated by the software. This even gives users the freedom to create their own pre-processing and diagnostic algorithms to create diagnosis models from scratch. This enables users to capitalize on their own knowledge in data analysis when creating a diagnosis model.

MELSOFT MaiLab provides a function block to embed python code written by users as well. This lets users link MELSOFT MaiLab with their own original processes to improve the accuracy of the diagnosis models (Fig. 7).

3.3 Use of Data for Greater Acceptability

On-site application of diagnosis models for data-driven improvements to production facilities must achieve everything from status monitoring to predictive detection using the data generated in real time on site. Typical data analysis software does not generally offer a function to use diagnosis models in production operations. This usually requires users to configure a separate system. MELSOFT MaiLab can both create diagnosis models and facilitate use of those models in production either on the same computer or deployed to another computer for use in production as is (Fig. 8).

Additionally, MELSOFT MaiLab provides functions to link data with the Mitsubishi Electric MELSEC-series programmable controllers and basic Edgecross software. These links can shorten the time required from implementation to use in production as well as reduce costs.



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

Fig. 6 Model editing operations in the editor

Pre-processing Function Block					
	Imputation process		Outlier handling process		Scaling
	Numerical operation		Dimensionality reduction		Set generation (Waveform extraction)
	Set feature generation		Subset formation		
Analytical function block					
	Deep learning		Gradient boosting decision tree		Randomized forest
	Auto encoding		k-nearest neighbors algorithm		Similar waveform recognition
	Multiple regression		MT method		Guard band

MT: Mahalanobis−Taguchi

Fig. 7 List of analysis function blocks

4. Application Example

MELSOFT MaiLab facilitates the creation of diagnosis models and use of those models on-site in production, even without a data scientist, by leveraging deep learning and other machine learning methods. As one example of issues to data-driven production on-site, many manufacturing businesses face the challenge faced when trying to share and utilize the knowledge of skilled workers.

4.1 Recommendations for optimal control parameters

Skilled workers can make decisions on the best settings for control parameters based on their intuition and experience. The problem is that this kind of knowledge is tacit knowledge, which is lost when those skilled workers retire.

MELSOFT MaiLab sets and learns from production data and the control parameters decided upon by these skilled workers to create diagnosis models that propose control parameter settings equivalent to those set by skilled workers (Fig. 9). In this way, the software enables employees with little experience to set the same parameters as skilled workers.

4.2 Estimation of the probable causes of problems

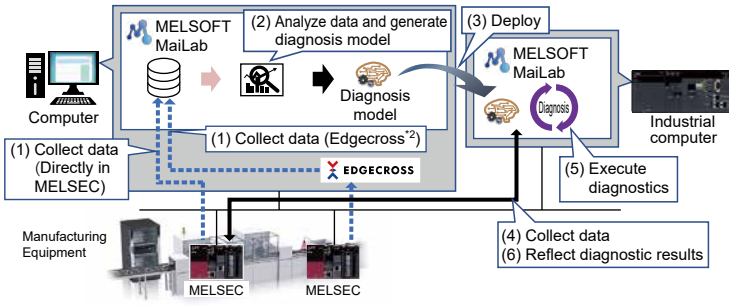
Anyone can estimate the probable causes of device malfunctions or product defects by investigating the causes of a past issue. However, it requires time for an employee with little expertise and experience to directly tackle these types of problems.

MELSOFT MaiLab can infer the probable causes based on the status when a problem occurs by learning from data about the status and type of issues encountered in the past. These capabilities support investigations into the cause of problems that require experience and expertise, which in turn shortens the time required to identify the probable causes. As a result, manufacturers can reduce the downtime of equipment and help improve productivity.

5. Conclusion

In light of the business landscape, Mitsubishi Electric developed MELSOFT MaiLab as a software with a variety of features to offer total support of data use on production sites.

Data analysis and AI are rapidly evolving. As these powerful technologies continue to become more diverse, the industry faces a serious lack of human resources and skills in data analysis. Mitsubishi Electric will continue to contribute to manufacturing solutions and production enhancements by broadening new technologies and expanding the application scope.



*2 Edgecross is a registered trademark of the Edgecross Consortium

Fig. 8 Example of the system configuration

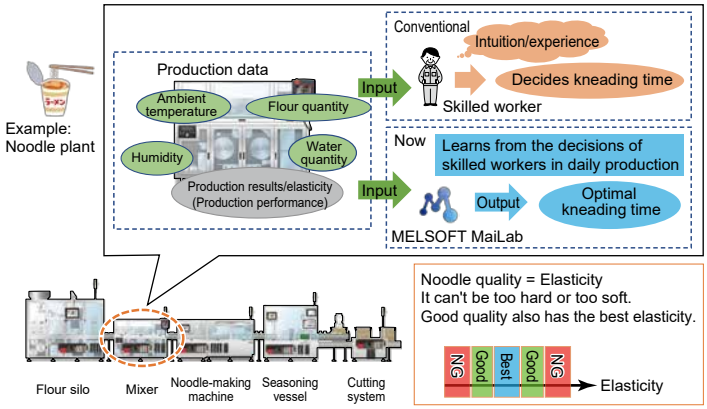


Fig. 9 User case to propose optimal control parameters