Data Analytics as Core of Value Creation

IIZAKA, Tatsuya[*]   DAITO, Takeshi[*]   MATSUI, Tetsuro[*]

ABSTRACT

The advance of the Internet of Things (IoT) facilitates the collection of a large amount of data. Data analytics is the core technology for creating customer value. There are 3 major challenges to apply this technology in the industrial context: (i) impractical accuracy when an insufficient number of cases are available for learning, (ii) limited application to areas that require clear explainability of results, and (iii) complex modeling for ordinary system designers. To solve these issues, Fuji Electric has leveraged its original innovations and developed data analytics tools. The implemented analytical methods ensure wide applicability, with a particular focus on diagnosis and forecasting, which are widely demanded.

1. Introduction

Recently, it has become possible to easily collect a large amount of data thanks to the development of the Internet of things (IoT). This creates solutions that produce various values one after another. This paper explains data analytics, which is the core technology for creating solutions and serves as the core of value creation.

2. Challenges to Be Solved by Data Analytics

The data analytics of Fuji Electric is a general term for statistics, machine learning and artificial intelligence technologies to perform recognition, diagnosis, prediction and optimization (see Fig. 1). These technologies progress quickly worldwide. In some fields such as image recognition and games like go and shogi (Japanese chess), they outperform humans. However, in the industrial field particularly many clients need solutions. There are the 3 following major challenges of recognition, diagnosis and prediction, and various efforts are being made for them.

(1) Challenge a: Practical accuracy cannot obtained from an insufficient number of learning cases.

In the industrial field, terms and data tendencies are different depending on the application. Therefore, there is small number of learning cases, and a practical level of accuracy cannot be obtained.

(2) Challenge b: Inference base cannot be explained

Since the basis for the inference result is unknown, and the application is limited in a field that requires reliability.

(3) Challenge c: Complex modeling for ordinary sys-

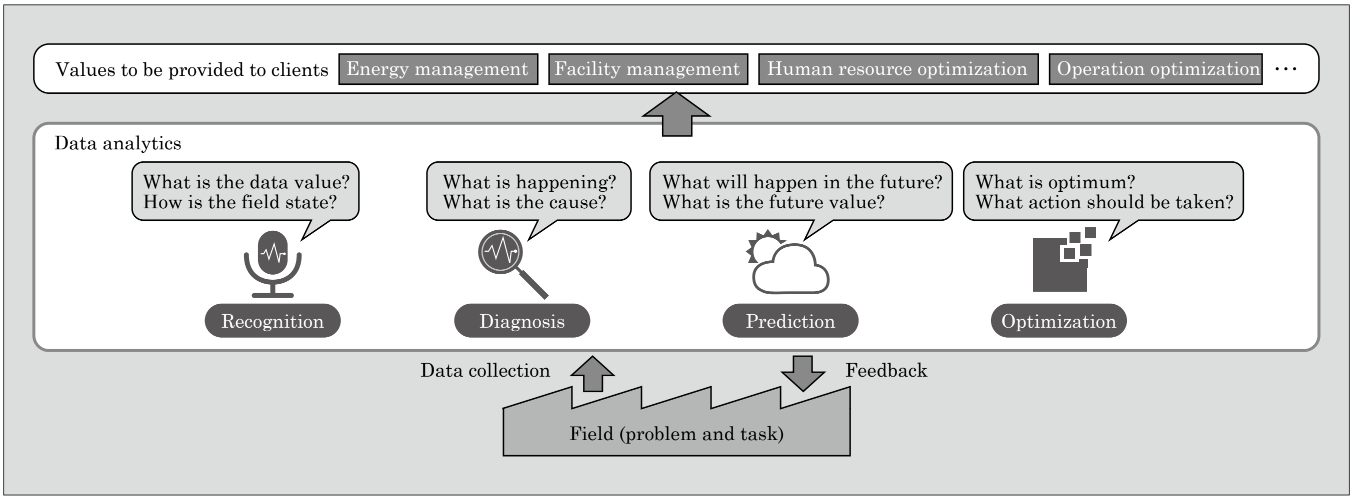

Fig.1 Data analytics

* Corporate R&D Headquarters, Fuji Electric Co., Ltd.

tem designers

To construct high-performance diagnosis and prediction models, trial and error learning by data scientists is necessary, and it is difficult for general designers.

3. Whole Image of Data Analytics

Solutions using our data analytics are classified into recognition, diagnosis, prediction and optimization (see Fig. 1). For recognition, solutions such as voice input, image recognition, and semantic analysis improve efficiency of maintenance work at a manufacturing site to create customer values. For diagnosis, solutions such as abnormality diagnosis and remaining lifespan estimation, improve the yield rate and quality to create customer values. For prediction, solutions such as demand prediction and sign detection reduce cost and supporting driving to create customer values. For optimization, solutions such as optimal operation and power generation planning for energy plants reduce fuel cost to create customer values.

4. Introduction of Data Analytics

Data analytics is a core technology that can realize various solutions by learning a large amount of data, constructing high-precision models and utilizing them. This chapter explains our unique technology that can solve the challenges in Chapter 2.

4.1 Text recognition technology

Text recognition technology generally consists of 3 types of processing.

- ○ Dividing sentences into words and parts of speech
- ○ Converting synonyms into the same words
- ○ Processing, extracting and aggregating sentences in accordance with the purpose of analysis

The important processing part of text recognition technology is converting synonyms into the same word. Technology for automatically creating synonyms from a large number of documents was developed, and the accuracy is high for general terms where it is easy to collect learning cases. However, in the industrial field where learning cases are limited, there is problem of impractical accuracy, as described in Challenge *a* of Chapter 2.

To convert synonyms into the same word, we create 2 types of dictionaries, one for general terms learned with a large amount of cases and one for technical terms learned with limited cases. The result of each dictionary's sentence processing is integrated (see Fig. 2). Following these approaches[1], we have realized a high recognition accuracy by using general terms to complement the industrial fields where there is a difficulty in collecting learning cases.

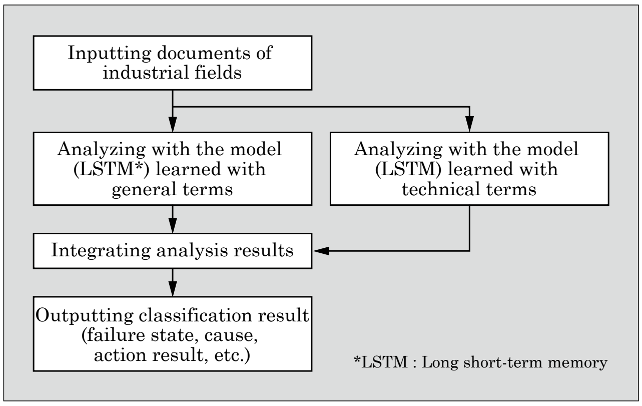

Fig.2 Text analysis method using several dictionaries

We are currently developing a plant maintenance system using this text recognition technology. The system can shorten the failure time by extracting records (failure state, cause and action result) that are close to the current state of the plant from the past maintenance records.

4.2 Diagnosis technology

We have applied multivariate statistical process control (MSPC), which has been used in the chemical process field, to diagnosis technologies such as abnormality diagnosis. We are developing abnormality diagnosis technology using machine learning to expand the application to objects having more complicated characteristics.

There are various diagnosis methods for complicated characteristics. In this section, the kernel-principal component analysis (PCA) method will be used as an example.

In normal machine learning, it is necessary to have learning cases of both normal data and abnormal data. However, there are few learning cases of abnormal data in the actual manufacturing process, and no practical accuracy could be obtained (Challenge *a*). The kernel-PCA solves this problem by performing learning only with abundant normal data and detecting states that are different from normal data (abnormality).

The kernel-PCA enables diagnosis even with non-linear characteristics by mapping data to a higher dimensional space using kernel functions and performing PCA in the higher dimensional space. Normal kernel-PCA can diagnose normal and abnormal states but does not have methods for determining the cause (vibration, voltage, etc.) when it diagnoses the state as an abnormality (Challenge *b*). Thus, we expanded advanced technology called the reconstruction based contribution (RBC) method[2] and applied that to kernel-PCA so a diagnosis basis similar to a normal MSPC can be presented (see Fig. 3). As a result, it is possible to quickly cope with an abnormality in the client equipment.

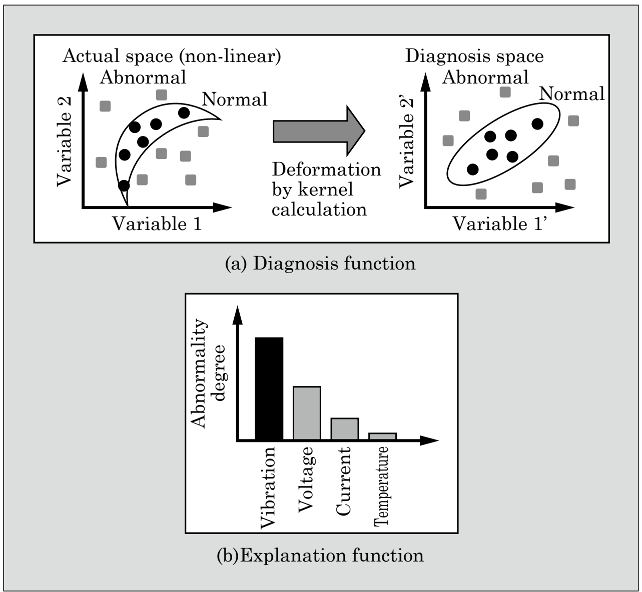

Fig.3 Diagnosis method using kernel-PCA

4.3 Prediction technology

Fuji Electric is developing various prediction technologies for objects. Just in time (JIT) modeling prediction is mainly used for objects having a limited amount of learning data, and a hierarchical neutral network such as deep learning is used for objects that can obtain a large amount of learning data*1.

JIT prediction is a technology that extracts data similar to the prediction object from the past data to make a prediction (see Fig. 4). With respect to Challenge *a* and Challenge *b*, the power demand and the merchandise sale amount are predicted as target objects for example based on similar cases. Therefore, a practical level of prediction is possible with objects with limited learning data, and the basis for the prediction is easy to understand (it is possible to present the basis for a prediction such as "the temperature is similar to that of the prediction target day" and "the day of the week is the same"). On the other hand, regarding Challenge *c*, developers needed to create models that define the definition of "similar" through trial and error. Regarding objects with changing consumption trends such as power demand and merchandise sales amounts, the definition of "similar" needs to be updated on a daily basis. To deal with this problem, we have used a decision tree, one type of machine learning, and developed a method that automatically defines the degree of "similarity." This method analyzes the recent trend of the prediction object with a decision tree, and automatically defines the similarity according to the variable importance obtained from the decision tree.

This method has been adopted in "ECONO-CREA," a power company cloud system and a solution for new power where there is a limited learning data amount because of the recent electricity deregulation. The system has been jointly developed by Fuji Electric, NTT DATA Corporation, and Kyowa Exeo Corporation. It won the Minister of Internal Affairs and Communications Award and general grand prize of ASP and SaaS section at ASPIC IOT Cloud Award 2017.

To cope with both Challenge *b* and Challenge *c*, we have added our unique approach to the hierarchical neural network. Regarding Challenge *b*, we designed the network structure, and extracted and visualized the correlation between input and output to indicate the basis for prediction. Regarding Challenge *c*, neurons counts and the combined state were conventionally determined by trial and error. The method we developed determines whether the neurons in the hidden layer is working effectively or unnecessary, and can automatically delete unnecessary neurons in the hidden layer. Figure 5 shows the analysis flow when a 4-layer network learns the data obtained by mixing 8 kinds of function data having different characteristics. It can be explained by extracting the characteristics of the input and output learned in the network in a function shape. The 3-layer structure is being applied to demand prediction of large electric power companies, prediction of streamflow into dam, and remaining life assessment of transformers. Regarding a network structure with 4 or more layers, the principle develop-

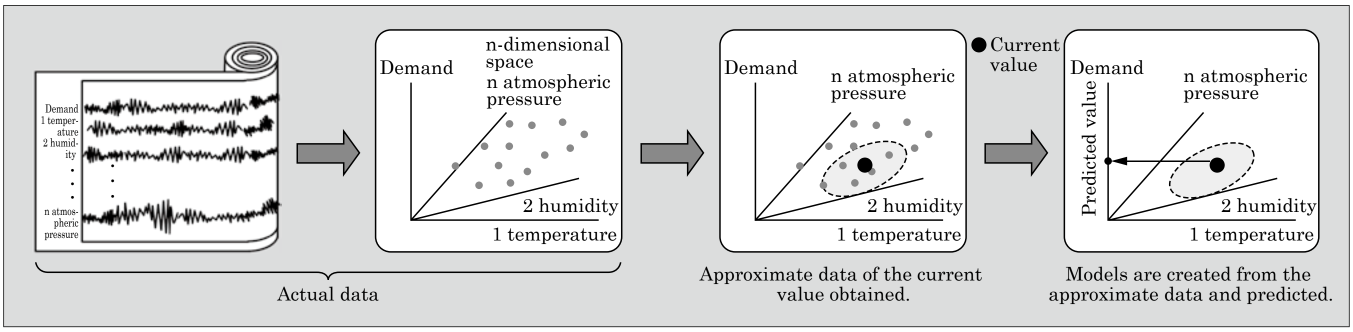

Fig.4 Conceptual diagram of JIT prediction

*1: Neutral network, Deep learning: Refer to "Supplemental explanation 1" on page 158.

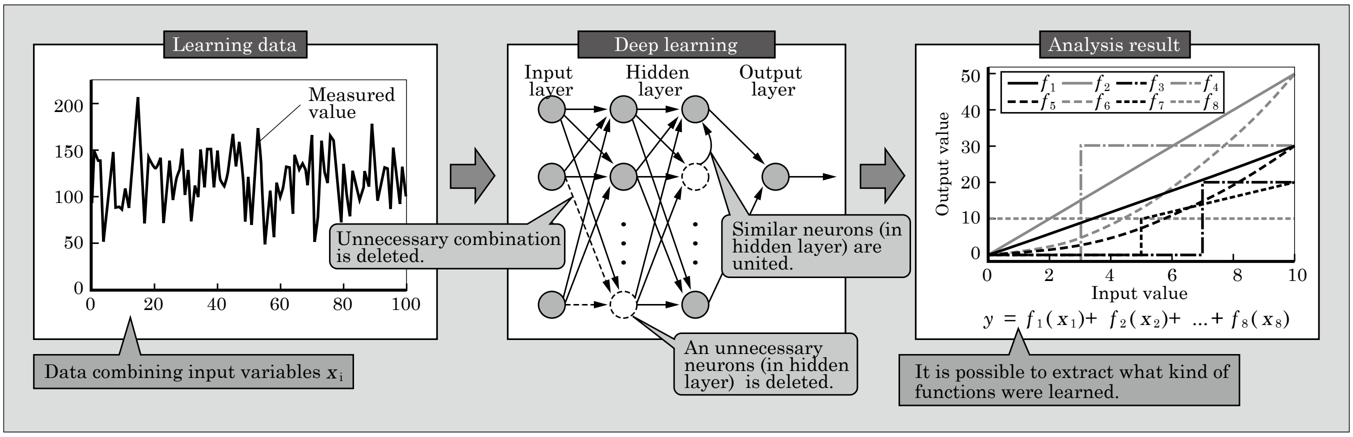

Fig.5 Analysis flow of deep learning

ment has been done, and the operation support system of the plant is being intensively developed.

5. Data Analytics Tool

Various types of data can now be easily collected by IoT; however, in contrast, there are not enough data scientists to analyze the large amount of data. Therefore, we are developing data analytics tools so other people besides data scientists can easily analyze data. This tool is particularly directed toward diagnosis and prediction, which are in high demand. Several methods are mounted so it can be applied to various targets (see Table 1 and Fig. 6).

The common function would be a cleansing function. The cleansing function can easily remove, replace, and complement the missing values and abnormal values as data pre-processing that generally takes up 80% of data analysis. Thus, the time needed for data analysis can be shortened.

Table 1 Function outline of data analytics

Category	Function	Method
Common	Cleansing	Removal, replacement and complement of missing values and abnormal values
Diagnosis	Visualization	Principal component analysis (PCA)
		Generative topographic mapping (GTM)
		t-distributed Stochastic Neighbor Embedding (t-SNE)
	Modeling	PCA
		Partial least squares (PLS)
		Kernel-PCA
		Support vector machine (SVM)
		eXtreme Gradient Boosting (XGBoost)
Prediction	Modeling	PLS
		Just in time modeling (JIT)
		Hierarchical neural network (deep learning)
	Quality simulation	PLS
		JIT

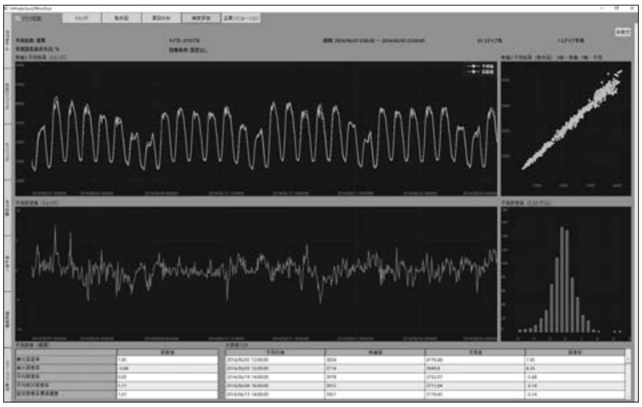

Fig.6 Example of prediction screen using data analytics tool

The visualization function in diagnosis determines the possibility of diagnosis. It compresses the multidimensional data into 2-dimensional data, and determines the data to be diagnosable if the abnormal data and the normal data are separately shown, and not diagnosable if the data cannot be separated. This function can determine whether the target data is diagnosable.

Figure 6 shows an example prediction screen of a data analytics tool. The reason for implementing several methods is because suitable technologies are different depending on the characteristic complexity of the diagnosis and prediction target. It is not easy for data scientists to determine which method is the best, and it has conventionally been determined by trial and error. Therefore, the tool has a function of automatically presenting the best method. A method that can handle complex characteristics also learns noise, and the accuracy may be reduced in the case of actual use. To prevent this, the method clearly separates the learning data and the test data so verification can be done under the same conditions as in actual use. This mechanism allows those who are not data analysts to easily analyze data and verify the results.

The quality simulation function determines the best manufacturing conditions with the manufacturing factories and plants as objects. This function learns combinations of various normal conditions and quality data in the past as quality simulation models. By simulating a quality change when the manufacturing condition is changed, the manufacturing condition that improves the yield rate and the quality can easily be obtained.

6. Application Example

The following refers to an example of applying the system to the temperature prediction of a manufacturing plant using deep learning. It predicts the temperature of intermediate products of several hours later from several tens of types of measurement data such as fuel and raw material quality. The temperature of intermediate products affects the quality of final products, and it is necessary to accurately control the temperature. Since the time constant of plants is long, it was necessary to predict the future temperature and control it in a feed-forward manner. The characteristic of the plant is not only complicated, but there is also a change in quality because of the difference in the lot or the origin of the raw material. Thus, it was difficult to automate with computers and controlling was done by human judgment. On the other hand, applying deep learning has made it possible to model complicated phenomena. To update the sequential model with close data, it became possible to deal with chronological change without human judgment.

As shown in Fig. 7, the change in the actual value can be properly predicted, contributing to the creation of customer values by improving the quality of products. Figure 8 shows an example of the analysis result of deep learning. When the ratio of a certain raw material component is increased, the temperature of the intermediate product drops. When the pressure rises, the temperature of the intermediate product rises. On the other hand, it is possible to quantitatively grasp that the flow rate hardly affects the temperature of intermediate products. That is, clients can use deep learning without anxiety by checking what kind of learning has been performed.

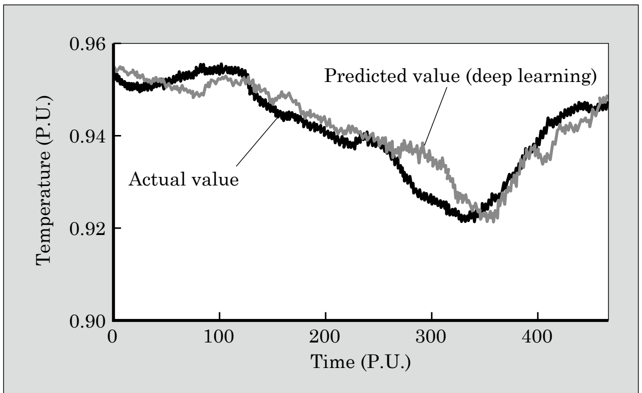

Fig.7 Temperature prediction result of manufacturing plant

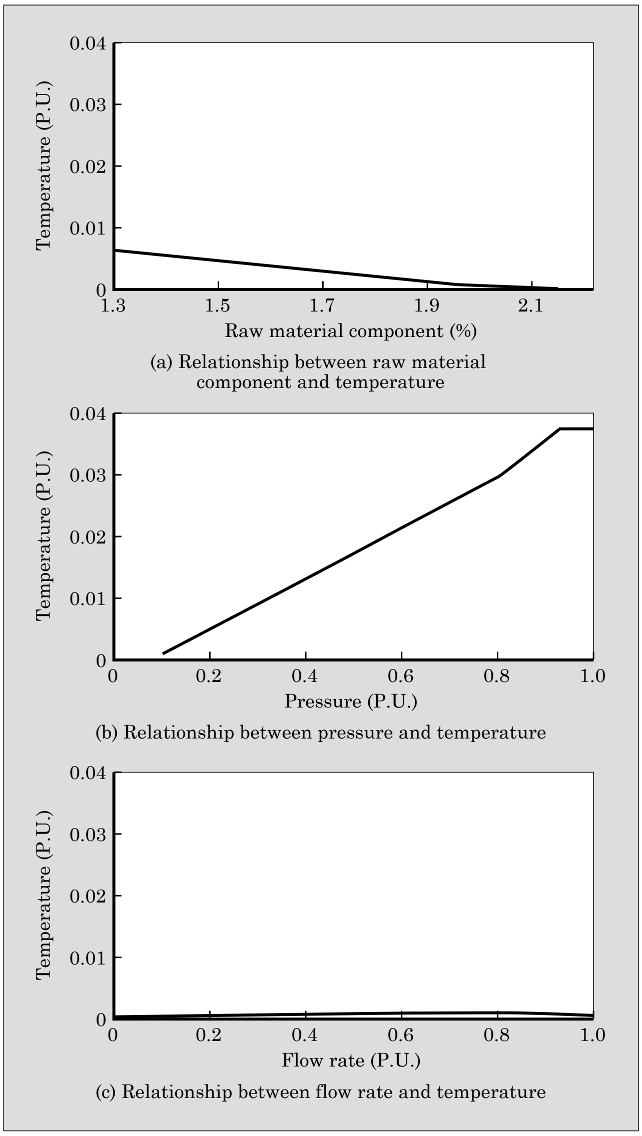

Fig.8 Example of analysis result using deep learning

7. Postscript

This paper describes data analytics, which serves as the core of value creation, has been described.

A large amount of data can be collected with IoT technology. However, conventional technologies could not handle an enormous amount of data sufficiently, and their performance result was not always satisfactory. To provide solutions that accelerate development using these huge data, we think that our unique data analytics is useful. Fuji Electric will expand the application to actual plants in the future to further create customer values.

References

(1) Yi Yang, Jacob Eisenstein. Unsupervised Multi-Domain Adaptation with Feature Embeddings. NAACL2015.

(2) Carlos F. Alcala, S. Joe Qin. "Reconstruction-based

Contribution for Process Monitoring with Kernel Principal Component Analysis". American Control Conference, 2010.

Fuji Electric
Innovating Energy Technology