Machine Learning Applied to Sensor Data Analysis: Part 2

Hiroaki Kanokogi [*1] *Go Takami* [*1]

In recent years, new technologies of machine learning and artificial intelligence (AI) have been progressing rapidly. Although there are high expectations for their application to industrial automation, there are various misunderstandings. For instance, some people incorrectly believe that AI will immediately deliver something new, but it is necessary to choose a suitable method to solve problems. Meanwhile, the pessimistic opinion that AI is eventually useless is also incorrect, as evidenced by impressive results in particular fields such as image recognition and board games. After all, what can machine learning actually do? This paper explains the characteristics of machine learning with examples of how Yokogawa uses it in plant data analysis.

INTRODUCTION

Expectations for machine learning and artificial intelligence technologies have increased in recent years and Yokogawa Electric Corporation and Yokogawa Solution Service Corporation have also released their specific cases of applying these technologies to Industrial Internet of Things (IIoT) [(1)(2)(3)].

Meanwhile, the reality of artificial intelligence is not well known to the public so there are many misunderstandings among people. Indeed, customer opinions are also mixed. Some have high expectations for the new technology but some are pessimistic.

- You can immediately obtain something new just by using artificial intelligence.
- Artificial intelligence is just the same as other analysis techniques (and will not be useful).

In conclusion, the above two opinions are both incorrect. Then, what is artificial intelligence or machine learning which is a part of artificial intelligence? This paper describes the features of machine learning applied to data analysis using specific examples and then presents Yokogawa's actual application cases for plant data analysis.

INTELLIGENCE OF MACHINE LEARNING AND DOMAIN KNOWLEDGE

Intelligence of Machine Learning

We use the artificial data example shown in Figure 1 to clearly explain the differences between the machine learning and conventional analysis methods.

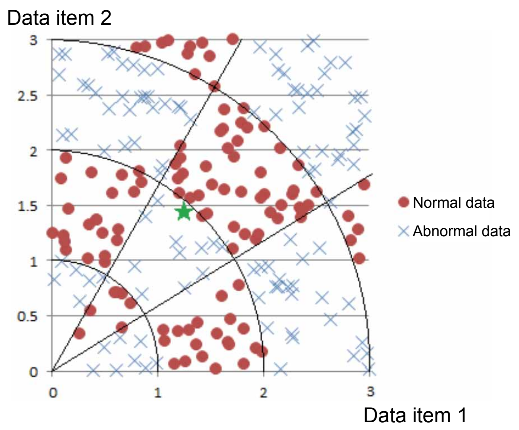

Figure 1 Data example of judging normal or abnormal

The normal or abnormal state is known for the horizontal axis (data item 1) and vertical axis (data item 2) data. Suppose you want to judge whether the green ★ data in the Figure is normal data (●) or abnormal data (×) (please think of this as a simple quiz to judge the ★ data only from information available in the Figure).

Intuitively, the most convincing answer would be that the star data in the graph of Figure 1 is × (abnormal data). Then, how about the following Figure 2?

---

*1 Field Digital Innovation Division, New Field Development Center, IA Product & Service Business Headquarters

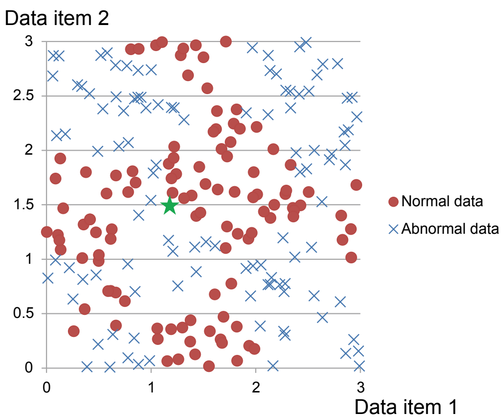

Figure 2 Data example 2 of judging normal or abnormal

Figure 2 is just Figure 1 minus the auxiliary lines. Therefore, the answer should be of course the same as with Figure 1 so the star data should be abnormal data (★=×). But compared with the case in Figure 1, you will feel that it is more difficult. To solve this problem correctly, you need to take the following three steps.

1. Look at the entire data and find regularities.
2. Draw auxiliary lines as simple as possible using straight lines or curves (equivalent to the auxiliary lines in Figure 1).
3. Regard the auxiliary lines as boundary lines and judge the ★ point.

Middle school and high school students may try to draw the lines in Figure 2 and may eventually find auxiliary lines in Figure 1. But it would be difficult for elementary school students to do so. In other words, the intelligence to find regularities is required to solve the problem in Figure 2 correctly.

Let's return to machine learning. If you put the data in Figure 2 into a machine learning program, will the program use its intelligence and take steps 1 to 3 as a human would? Figure 3 shows a typical judgment result by conventional analysis.

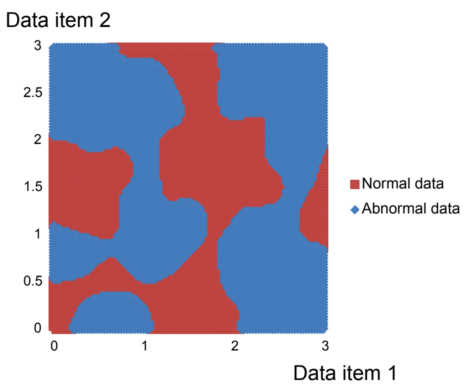

Figure 3 Judgment result by conventional analysis

If you compare Figure 2 and Figure 3, you will find that data near the known normal input data is judged as normal data and data near the known abnormal input data is judged as abnormal data. The judgement target ★ part happens to be near a cluster of normal data (●) so the data is judged wrongly as normal data. Obviously, there are no signs that a human-like search was made to find regularities by looking at the entire data.

Then, how does the judgment result by machine learning look?

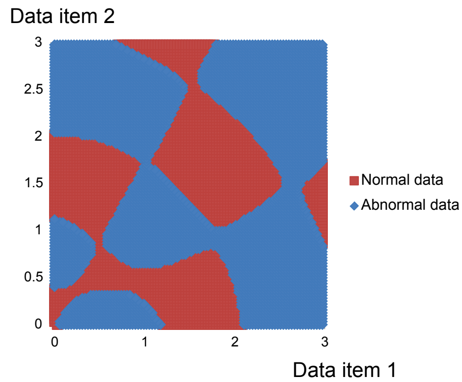

Figure 4 Judgment result by machine learning

In the result in Figure 4, you will find that the auxiliary lines seen in Figure 1 are approximately reproduced (divided by arcs in the radial direction and by straight lines in the deflection angle direction). This means that machine learning took steps 1 to 3 that would be taken by humans for the problem in Figure 2 (the ★ part was also judged correctly by the clear boundary line of arc).

As described above, a kind of intelligence is required even for humans to look at the entire data and draw auxiliary lines. Main machine learning techniques proposed in recent years have this kind of intelligence. The above example is a simple one, but machine learning performs the same kind of structure search even on 30-dimensional data that cannot be processed by humans. Machine learning is able to execute the capability possessed by humans to look at the entire data and simplify things by multivariable analysis of dimensions that cannot be processed by humans. Although it looks significantly different from the artificial intelligence that appears in science fiction, we think that it is one of the ways to use artificial intelligence for plant data analysis (note that there are many definitions for intelligence in artificial intelligence and the one described herein is one of them).

The ability to judge subtle normal and abnormal conditions as described in Figure 4 is particularly important for plant data analysis. What is really important in actual plant judgment results is in many cases the state diagnosis in the vicinity of the boundary line of normal and abnormal data, because, for example, an easy case where the judgment target ★ data is at the center of a normal data group does not need analysis. As is obvious from comparing Figure 3 and Figure 4, we think that the excellent panoramic and objective features of the machine learning techniques proposed in recent years are very important for accurate state diagnosis and preventive maintenance.

Meanwhile, Figure 4 shows that machine learning is not necessarily perfect. If the characteristic of being able to classify data by arcs and straight lines is known in the problem in Figure 2, using that domain knowledge (drawing known auxiliary lines) from the beginning makes the diagnosis significantly more accurate. For example, as for problems where the characteristics of data distribution can be predicted, such as those in quality engineering, using the Mahalanobis Taguchi (MT) method and conventional analysis techniques makes the accuracy higher.

The ability of machine learning described herein is called a generalization ability or statistical consistency and it is proved that it is a characteristic possessed by the main machine learning techniques proposed in recent years. If you have interest in mathematical theories, please refer to reference [(4)]. Furthermore, the description of fields in which conventional analysis methods (such as the MT method) and machine learning specialize, respectively, is based on reference [(5)].

No-free-lunch Theorem and Domain Knowledge

As described above, we think that artificial intelligence and machine learning are promising for plant data analysis. Meanwhile, the thought that "you can immediately obtain something new just by using artificial intelligence" is a widespread misunderstanding.

A theorem called a no-free-lunch theorem is well known in the learning theories.

> The no-free-lunch theorem [(6)(7)] is mathematical proof of the observation that "learning algorithms cannot be universally good for any problem."

In other words, "there is no universally applicable machine learning tool" or "an effort to use some domain knowledge is required to obtain good learning results." Every machine learning technique has its own characteristic. The no-free-lunch theorem is proof that if you do not use the characteristic suitable for the problem you have to solve, you cannot obtain the optimal solution. For example, to analyze sensor data, you must select the appropriate machine learning method using knowledge on what is noise dependent, what is the approximate ratio of the number of normal and abnormal data records for a sample, whether or not the judgment result can change statically or dynamically, and so on.

We ask customers about reasons why they select a particular machine learning method for plant data analysis. Many customers answer, for example, "because it is famous" or "the deployment price is affordable (or free of charge)." But if a machine learning method is deployed for such a reason, the initially expected results will not be obtained and the customers will regret it in many cases. According to our experience, there is almost no probability that a desired result can be obtained just by inputting plant data into a general-purpose machine learning tool (except for simple problems as described above). As with conventional analysis techniques, it is also important to use domain knowledge when using machine learning.

Domain knowledge is important even after analysis results are obtained. Plant data output of machine learning is basically ● or ×, or any numeric value. It is not a machine but a person that can add an interpretation to give more meaning than a numeric value.

ACTUAL EXAMPLE OF SENSOR DATA

The following describes an experiment at Yokogawa and an actual example at a customer's plant that applied machine learning to plant data analysis.

pH Meter Deterioration Diagnosis

Reference (1) shows the experimental results of pH meter deterioration diagnosis, pH sensor maintenance optimization, and preventive maintenance by machine learning. The following shows an additional experiment and its results.

We collected teaching data for learning over two months from May 2016 in the experimental environment shown in Figure 5. We placed two pH sensors, a new one (with no deterioration) and an old one (after deterioration), in a reagent solution and acquired sensor data for each one. Then, we acquired data for evaluation of a total of eight sensors, four new ones and four old ones over nine months from July.

The purpose of the experiment was to create a learning model using the teaching data acquired in the former period and to verify to what extent the deterioration of the eight pH sensors (new ones and old ones) can be identified in the latter period based on the model. The expected result was that a clear difference can be found between the new pH sensors and old ones in the diagnostic results based on the created learning model.

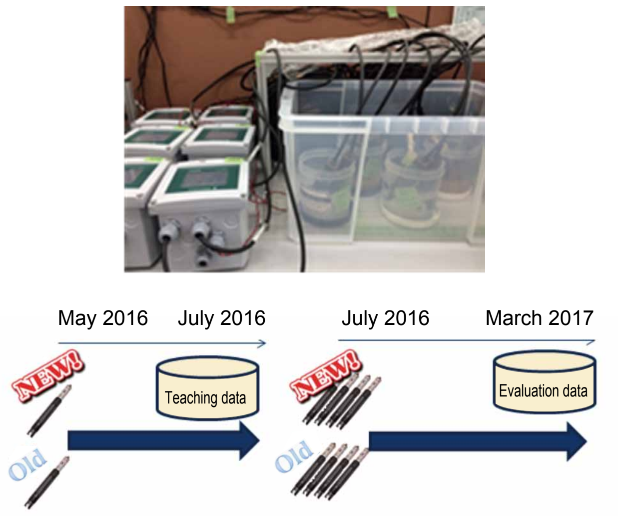

Figure 5 pH meter deterioration diagnosis experimental setup

First, we used a conventional analysis method (Fisher's linear discriminator) as an analysis technique in contrast to machine learning. Figure 6 shows the results obtained by creating a model by inputting teaching data into a program for this technique and applying the model to the evaluation data. The bar graph in the Figure shows the average value for deterioration judgment and the error bar shows variations in the judgment (standard deviations).

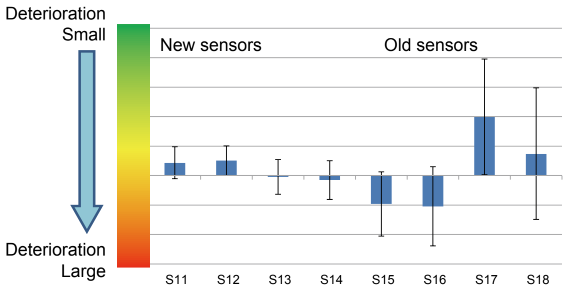

Figure 6 Deterioration diagnosis by conventional analysis

Old sensors S17 and S18 are judged wrongly as smaller deterioration than new sensors S11 to S14. Furthermore, the error bars for old sensors S15 and S16 are tall and deterioration is certainly difficult to judge as larger than for new sensors S11 to S14. Thus, with the conventional analysis method, the expected results were not obtained for the deterioration judgment of new sensors (S11 to S14) and old sensors (S15 to S18).

Meanwhile, Figure 7 shows the deterioration diagnostic results by machine learning.

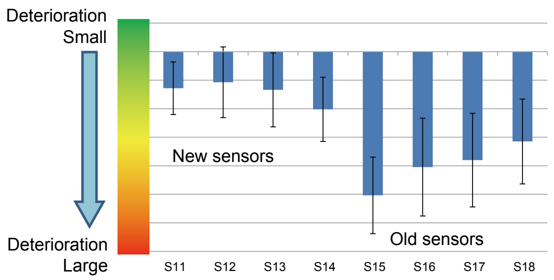

Figure 7 Deterioration diagnosis by machine learning

You can see that there is a clear difference between new sensors S11 to S14 and old sensors S15 to S18. There are no differences between the eight sensors except for new and old ones. Our opinion is that the differences in the judgment results in Figure 7 show that diagnosis by machine learning identified the deterioration phenomenon.

Compressor System State Diagnosis and Failure Sign Detection

The following presents an actual example of state diagnosis and failure sign detection of a compressor system at a plant using the machine learning tool that was used above. The following items 1 to 3 show an overview of the test target system.

1. Production system including a compressor
2. 25 types of sensor data, including pressure, flow rate, and temperature
3. Each sensor acquires data at one-minute intervals.

Figure 8 shows the diagnostic results of evaluation data in the unknown state using a model, which was created by inputting the normal and abnormal data of the above system into a machine learning program as with the pH deterioration diagnosis. You can see the internal state gradually shifting from normal to abnormal in (1) to (3). A large abnormality that actually occurred was detected in (6).

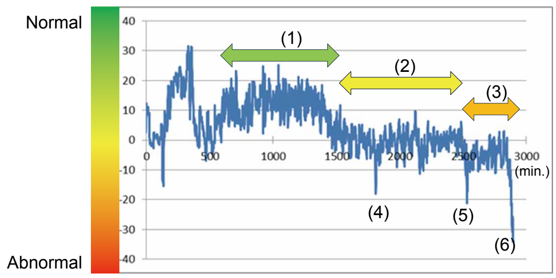

Figure 8 Compressor system abnormal signs

What is most important in the results of Figure 8 is detection of significant failure signs that appeared in (4) and (5). If you perform inspection at this point (of a sign in (4) about 16 hours before the sign (6) appears), you may be able to prevent a failure in (6) in some way (preventive maintenance). Thus, we believe that plant data analysis using machine learning is promising for state diagnosis and preventive maintenance. We compiled a final report that is composed of these calculation results and the interpretation and cause estimation using a wealth of domain knowledge owned by Yokogawa's specialized departments.

We also carried out the same analysis using the Fisher's linear discriminator described above and other methods but we could not obtain the signs shown in Figure 8.

CONCLUSION

We believe that machine learning specialized in the plant domain is very useful in plant's state diagnosis and preventive maintenance. We will expand the application range and work together with customers that provide support for analysis to accumulate real experience.

Finally, we would like to express our gratitude to Electric Power Development Co., Ltd. for support for the advanced research on the application of machine learning to plants and for approval to publish the results.

REFERENCES

(1) Go Takami, Moe Tokuoka, Hirotsugu Goto, Yuichi Nozaka, "Machine Learning Applied to Sensor Data Analysis," Yokogawa Technical Report English Edition, Vol. 59, No. 1, pp.27-30
(2) Tomohisa Shirai and Hideharu Nagasawa, "Innovation of Facility Management by Using IoT and Big Data," the 46th Annual Convention of the Japan Petroleum Institute (Kyoto convention), 2016 (in Japanese)
(3) Tomonori Sato, Souichiro Akiyama, Shinji Kawade, Shigekazu Kuroki, et al., "Machine Learning Aiming at Sensor Predictive Maintenance," the 60th Research Presentation Meeting of Kansai Division, Japan Water Works Association, 2017 (in Japanese)
(4) Takafumi Kanamori, Statistical Learning Theory, Kodansha, 2015 (in Japanese)
(5) Tsuyoshi Ide, Introductory Guide for Abnormality Detection by Machine Learning, Corona Publishing, 2015 (in Japanese)
(6) Wikipedia, No free lunch theorem, https://en.wikipedia.org/wiki/No_free_lunch_theorem
(7) Shinichi Asakawa, Deep Learning, Big Data, Machine Learning, or its Psychology, Shin-yo-sha, 2015 (in Japanese)