Wide Learning Technology to Provide Trust Through Knowledge Discovery

● Kotaro Ohori ● Tatsuya Asai ● Hiroaki Iwashita ● Keisuke Goto
● Junichi Shigezumi ● Takuya Takagi ● Yuri Nakao ● Hirokazu Anai

With the recent emergence of the AI black box problem, the fairness, accountability, and transparency (FAT) of AI have been vigorously debated. Companies and academic institutions have developed various types of explainable AI. Fujitsu Laboratories has developed a new technology, Wide Learning, which is not only capable of dealing with the black box problem but also includes explainability that achieves knowledge discovery. Wide Learning uses "enumeration," which is a major technology of discovery science, to exhaustively enumerate pieces of knowledge called knowledge chunks described in a human-understandable format. By using this knowledge for judgment, the overlooking of potential candidates is significantly reduced and high-accuracy prediction and classification are achieved. If knowledge without omission can be discovered continuously, the trust of service systems can be strengthened and knowledge diversion in turn should provide trust between service systems. This paper first outlines the history of how Wide Learning came into being and introduces its technical characteristics. It then reports on a demonstration of knowledge discovery using Wide Learning in Fujitsu and presents future prospects.

1. Introduction

In recent years, in the field of AI, a black box problem has arisen whereby it is not possible to explain the process whereby decisions are made in machine learning techniques such as Deep Learning. On December 27, 2018, the Cabinet Office of Japan held a review council for social principles of human-centric AI. At this council, a draft of seven principles of AI relating to social frameworks that should be implemented across Japanese society was published.[1] In addition to security and privacy, these include the principles of fairness, accountability, and transparency (FAT).

Also, in August 2016, the US Defense Advanced Research Projects Agency (DARPA) announced a research program on explainable AI (XAI).[2] Since then, businesses and academic institutions have been developing various forms of XAI, such as AI that ensures fairness by detecting data bias,[3] and AI that can explain the basis of a classification.[4]

Given this situation, Fujitsu Laboratories announced a new machine learning technology called Wide Learning in September 2018. The main difference between Wide Learning and conventional XAI is that Wide Learning can extract all the available knowledge from current data without omission, thereby greatly reducing the likelihood of overlooking items on which decisions should be made. From the viewpoint of trust, which is the theme of this issue, this is a role that creates trust in the ability to discover/acquire new knowledge from the huge quantities of data in service systems.

Also, when data is added, it is possible to notice changes in information at a glance, and to perform continuous knowledge discovery. Furthermore, by using the isomorphism of systems as a way of repurposing the resulting knowledge for other issues, we can expect to reduce the complexity of systems and promote new knowledge discovery.

In this way, Wide Learning technology has the potential to strengthen trust within a single service system as the core of knowledge discovery, with the resulting trust connecting to other service systems. This also overlaps with the system view of Society 5.0,[5] whereby multiple applications are connected together.

In this paper, we outline the history and background of the development of Wide Learning, and introduce its technical characteristics. We also describe its effects on service systems with reference to a number of examples of its use in internal Fujitsu projects.

2. History and background of Wide Learning

Progress in AI, especially machine learning, can be broadly divided into neural networks that mathematically model the brain's neural pathways, and symbol processing AI that is geared towards finding solutions by logical reasoning. Deep learning, which belongs in the same category as neural networks, has made remarkable progress in recent years and has achieved high prediction and classification accuracy. On the other hand, the latest technology for symbolic processing AI is Wide Learning, which is introduced in this paper.

Wide Learning emerged from a background of research into discovery science (DS), which was advocated by Professor Setsuo Arikawa in the 1990s.[6] Since DS was first proposed, Fujitsu Laboratories has been promoting research into knowledge discovery including data mining. We have also participated in numerous national projects and accumulated technical know-how in this field.[7),8)] In December 2015, we invited Professor Arikawa to establish the Arikawa Discovery Science Center at Fujitsu Laboratories. Since then, we have been working on theoretical and empirical studies of DS through open innovation together with outside experts.

Wide Learning is a new machine learning technique whose importance has become clear due to the accumulation of these research results. It involves the use of "enumeration" algorithms, which have evolved dramatically in recent years.

3. Overview and technical characteristics of Wide Learning

This section presents an outline of Wide Learning and discusses its technical characteristics.

3.1 Overview of Wide Learning

Wide Learning is a technique for enumerating knowledge fast that can be directly understood by humans (knowledge chunks) from combinations of data items, and using this knowledge for predictions and classifications (**Figure 1**). In Wide Learning, a fast enumeration technique is used to exhaustively search all combinations of data items, and all those whose appearance rates differ significantly between positive and negative examples are extracted as knowledge chunks.

An important elemental technology of Wide Learning is enumeration, which seeks to obtain an overall picture of problems by exhaustively listing their possibilities. Even if the conditions are intricately entangled and difficult to formulate, humans will still be able to trust the solution if it is possible to enumerate all the specific cases that constitute the problem. For example, consider the crane-turtle problem shown in **Figure 2**. This problem is easy to solve if you know how to do simultaneous equations. However, even younger children who do not understand such abstract concepts should still be able to understand the answer if all the

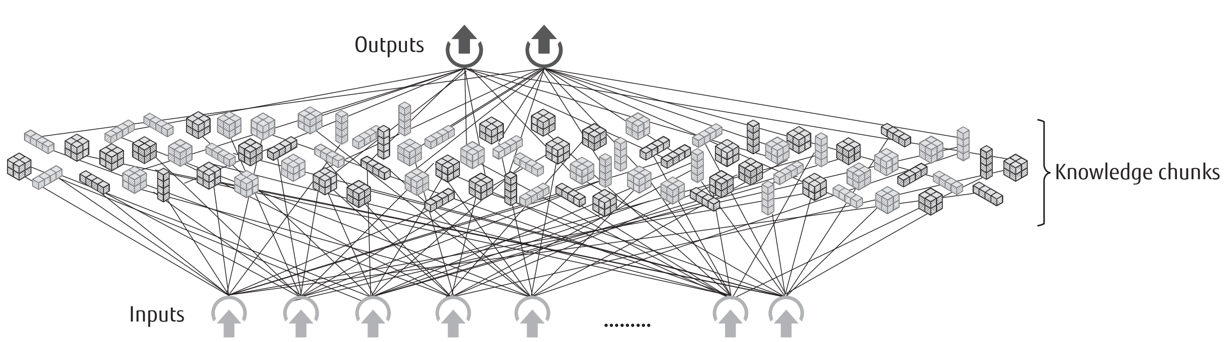
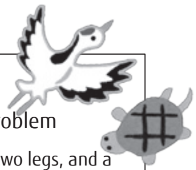


Figure 1
Overview of Wide Learning.

possibilities are listed.

Wide Learning makes it possible to gain a deep understanding of the characteristics of service systems through repeated knowledge discovery based on enumeration, and can strengthen people's trust in system designers and beneficiaries. In addition, an understanding of the characteristics of a specific service system can provide design guidelines for related systems, which can be expected to result in a chain of trust (**Figure 3**).

The following subsections describe the discovery and repurposing of knowledge, which are the technical characteristics of Wide Learning.



Example problem

A crane has two legs, and a turtle has four.
There are cranes and turtles, ten in total.
All together, 28 legs.
How many cranes are there, and how many turtles?

Example of enumeration

Cranes	Turtles	Legs
1	9	38
2	8	36
3	7	34
4	6	32
5	5	30
6	4	28
7	3	26
8	2	24
9	1	22

Correct answer

Figure 2
A crane-turtle problem and its enumeration

3.2 Fast enumeration techniques for knowledge discovery

As one of the core technologies of DS, enumeration is clearly an area where computers are better than humans. However, due to the phenomenon of combinatorial explosion, there are limits to the problems that can be solved by simple methods, even with a computer.

Suppose you want to identify the features of a product's customer base by combining criteria related to customer attributes, such as whether they are male or female, whether or not they have a driver's license, and whether they are married or single (e.g., "male, married"). With regard to gender, there are three possible outcomes: restricted to men, restricted to women,

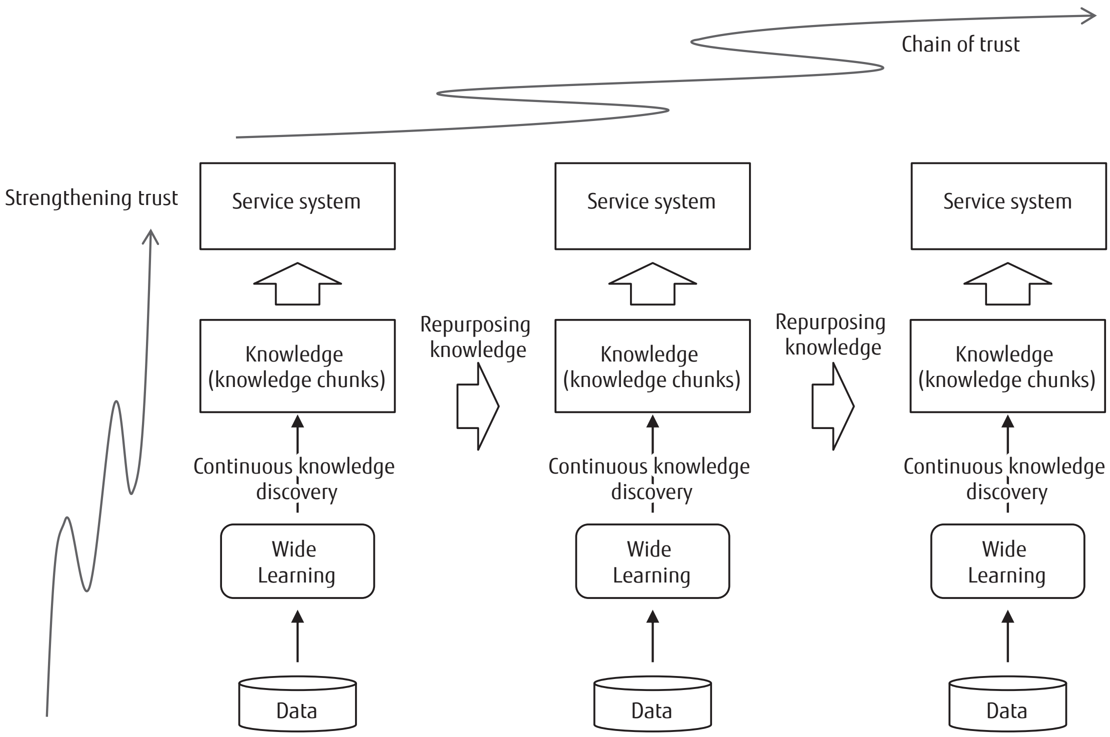


Figure 3
Connecting trust in Wide Learning.

and not restricted by gender. If there are ten attributes of this nature, then there will be 3^{10} (i.e., 59,049) possible combinations. At this sort of scale, it should be possible to use a computer to exhaustively check each combination properly represents the customer base. However, the number of combinations rises sharply with the number of attributes. For 20 attributes, there are about 3.5 billion combinations, and with 30 attributes there are over 200 trillion combinations. Thus, in this sort of simple enumeration, it is soon necessary to perform enumeration more efficiently to overcome the combinatorial explosion.

Two techniques that aid efficient enumeration are pruning and memoization. Pruning is a technique used when programming computers to play games like chess, and involves excluding positions that are clearly disadvantageous when analyzing possible moves. On the other hand, memoization is a technique for storing the analysis results for each set of conditions, and re-using these results later on to avoid calculating the same results later on. Another technique that is similar to memoization avoids duplicated calculations by performing subsequent analysis using procedures that are narrowed down to arrive at the same set of conditions.[7),8)] In Wide Learning, knowledge chunks for real-scale problems are enumerated fast by accumulating these techniques and optimizing the processing algorithms to maximize their efficacy.

Wide Learning performs functions such as classification and prediction by learning the weight and importance of the enumerated knowledge chunks. This was actually performed in an experiment using heart disease (medical) and bank marketing (financial) data from the UC Irvine Machine Learning Repository.[9)] As a result, it was confirmed that Wide Learning yields a correct answer rate roughly 10–20% higher than that of Deep Learning, while reducing the probability of missing correct answers by roughly 20–50%.[10)] As mentioned above, Wide Learning works in a different way to Deep Learning, so it is difficult to make a direct comparison. Nevertheless, this result seems to suggest that Wide Learning is an effective approach to exhaustive enumeration of knowledge that can reduce the possibility of overlooking potential candidates.

3.3 Expectations for knowledge repurposing

Many machine learning methods based on Deep Learning and existing symbol processing AI techniques are based on the concept of "optimization." Even if there are multiple models that achieve almost the same prediction accuracy, only one of them is adopted for optimization. Therefore, the prediction model may change greatly even if the input changes only slightly, or even if the optimization is simply re-run. This phenomenon makes it difficult to compare and repurpose learning results.

On the other hand, in Wide Learning based on enumeration, every useful knowledge chunk is adopted without omission. Therefore, not only are the same set of knowledge chunks always obtained from the same input, but it is also possible to obtain a set of knowledge chunks with many points in common from an input with many points in common.

In many fields, efforts are being made to speed up the preparation of environments for AI applications. On the other hand, there are still many fields where not enough data has been accumulated for AI applications. If AI is applied while available data is insufficient, its performance may not improve as expected, or it may exhibit unexpected behavior when presented with unknown data, resulting in poor reliability. However, it is expected that Wide Learning's strict knowledge extraction capabilities can be used to supplement a model that has insufficient data by repurposing universal knowledge discovered from a similar system that has plenty of data.

4. Wide Learning use cases

This section introduces an example of Wide Learning applied to knowledge discovery using marketing case studies. The possibility of repurposing knowledge by Wide Learning is also discussed.

In this example, we applied Wide Learning to customer data held by experts in Fujitsu's global marketing division, and we attempted to perform knowledge discovery and extract important features (knowledge chunks) of prospective customers for trade negotiation X. The customer data includes logs of their behavior such as the seminars and websites they visit, as well as attribute information such as their industry and job title. To keep the explanation simple, the number of constituent items in the knowledge chunks to be extracted is limited to two. We applied Wide Learning to customer data consisting of 300 features and extracted

Table 1
Expert opinions of extracted knowledge chunks.

Knowledge chunk judged to be important by Wide Learning	Expert opinion
[Job position: Subsection manager/chief class]	It is not only managerial posts such as managers and section managers that have a large influence, but also site leaders
[Interest in service: Large] and [Access to other users in the same company]	Initiating service actions not only by people but also by business units
[Interest in other services: Large] and [Access within 12 weeks]	When trade negotiations are approaching, check various other Fujitsu services in addition to the service being negotiated

1,300 importance-weighted knowledge chunks from a combination of approximately 180,000 data items with two or fewer items. Furthermore, we presented these weighted knowledge chunks to experts, who considered its validity and background. **Table 1** shows the knowledge chunks that were judged to be important in Wide Learning and a number of examples of their opinions.

The expert opinions included favorable comments. Not only were the Wide Learning results deemed to be "empirically correct," but we also received comments such as "I have never noticed this before, but it is a possible important feature." The series of processes performed in Wide Learning is based on data and does not involve any human beliefs or experiences. For this reason, it is highly likely that knowledge chunks acquired by Wide Learning may include knowledge that would not have occurred to humans. It is thought that this ability to bring latent knowledge to the surface will lead to the discovery of new knowledge.

Furthermore, the knowledge chunks shown in Table 1 consist of important features that are likely to appear in other business negotiations. When trialing this technique, we were informed that it managed to extract universal knowledge that was not only applicable to trade negotiation X, but also to other trade negotiations. In other words, if another trade negotiation Y does not have a sufficient data set, it should be possible to rapidly increase the accuracy of these negotiations by repurposing the knowledge chunks obtained in trade negotiation X.

5. Conclusion

This paper described Wide Learning—an explainable AI technique based on the enumeration of knowledge chunks. This technique not only enables precise classification, but also supports knowledge discovery and knowledge repurposing.

We are currently conducting trials of this technique, primarily in the digital marketing field, but in the future, we plan to expand the scope of application to other fields such as manufacturing and medicine. We also plan to speed up the development of techniques for accelerating the repurposing of knowledge, and to implement cooperation between Wide Learning and service systems via knowledge chunks.

References

1) Japanese Cabinet Office: Social Principles of Human-Centric AI.
https://www.cas.go.jp/jp/seisaku/jinkouchinou/pdf/humancentricai.pdf
2) M. Turek: Explainable Artificial Intelligence (XAI). DARPA.
https://www.darpa.mil/program/explainable-artificial-intelligence
3) A. Mojsilovic: Trust and Transparency for AI on the IBM Cloud. IBM.
https://www.ibm.com/blogs/research/2018/09/trust-transparency/
4) E. Webber: Similarity Cracks the Code Of Explainable AI. simMachines.
https://simmachines.com/similarity-cracks-code-explainable-ai/
5) Japanese Cabinet Office: Society 5.0.
https://www8.cao.go.jp/cstp/english/society5_0/index.html
6) S. Arikawa: Discovery Science Revisited. FUJITSU Sci. Tech. J., Vol. 53, No. 5, pp. 3–7 (2017).
https://www.fujitsu.com/global/documents/about/resources/publications/fstj/archives/vol53-5/paper01.pdf
7) H. Iwashita et al.: Efficient Computation of the Number of Paths in a Grid Graph with Minimal Perfect Hash Functions. Technical Report, Graduate School of Information Science and Technology, Hokkaido University, 2013.

http://www-alg.ist.hokudai.ac.jp/~thomas/TCSTR/tcstr_13_69/tcstr_13_69.pdf
8) ERATO Minato Discrete Structure Manipulation System Project: Ultrafast graph enumeration algorithm. Morikita Publishing, 2015. (in Japanese)
9) P. Murphy et al.: UCI repository of machine learning databases. Department of Information and Computer Science, University of California, Irvine, California, 1992.
10) Fujitsu Laboratories: Fujitsu's New AI Technology "Wide Learning" Enables Highly Precise Learning Even From Imbalanced Data Sets.
https://www.fujitsu.com/global/about/resources/news/press-releases/2018/0919-03.html

Kotaro Ohori
Fujitsu Laboratories Ltd.
Dr. Ohori is currently engaged in research of social system designs using AI and mathematical technologies.

Tatsuya Asai
Fujitsu Laboratories Ltd.
Dr. Asai is currently engaged in research of AI based on discovery science.

Hiroaki Iwashita
Fujitsu Laboratories Ltd.
Dr. Iwashita is currently engaged in research and development of knowledge discovery technology based on discrete mathematics.

Keisuke Goto
Fujitsu Laboratories Ltd.
Dr. Goto is currently engaged in research and development of string processing and AI.

Junichi Shigezumi
Fujitsu Laboratories Ltd.
Dr. Shigezumi is currently engaged in research of AI related to knowledge processing, information retrieval, inference, etc.

Takuya Takagi
Fujitsu Laboratories Ltd.
Dr. Takagi is currently engaged in research and development of string processing, data mining, and AI.

Yuri Nakao
Fujitsu Laboratories Ltd.
Mr. Nakao is currently engaged in research of recommendation systems, fairness-based machine learning, and user observation of accessibility.

Hirokazu Anai
Fujitsu Laboratories Ltd.
Dr. Anai is currently engaged in research and development of the theory and application of mathematical processing, optimization, and AI.