

FUJITSU Project:

Title Enhancing explainability of modern AI

Industrial Partner FUJITSU LIMITED

Fujitsu, a leading Japanese information and communication technology (ICT) company, offers a full range of technology-related products, solutions, and services. We have produced numerous and diverse innovative results through IT and have been leading the world by providing new value to individuals, society, and businesses. Fujitsu, number 1 in Japan, is the world's tenth largest IT service provider.

Industrial Mentor

Hiroyuki Higuchi, Ph.D., Fujitsu Limited.

Background

Artificial intelligence (AI) has been increasingly introduced into various industries. Nevertheless, in many cases, effective utilization is not achieved, mainly because AI provides users no information related to its judgments. Therefore, a strong need exists for modern and practical AI to offer explainability to the users and present them reasons to justify the AI rationale [1].

Fujitsu has developed “Wide Learning”: explainable AI (XAI) that combines discovery science and machine learning [2,3]. Figure 1 depicts the principle of Wide Learning with an example of marketing data. An operator enters multiple aspects of data of a customer who has purchased a product. Wide Learning exhaustively extracts important combinations of data items as hypotheses hidden in various data and determines the likelihood of new customers purchasing a product. The important hypotheses are explained as reasons supporting the decision.

Table 1 presents a comparison between Wide Learning and Deep Learning [4].

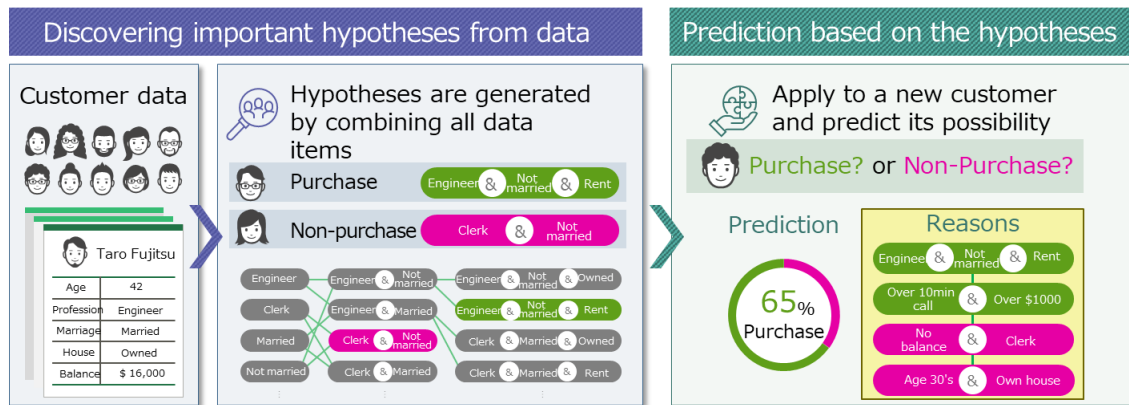


Fig. 1 Principle of Wide Learning.

Table 1 Wide Learning and Deep Learning

	Wide Learning	Deep Learning
Operating principle	Process of scientific discovery	Simulation of a neural network
Suitable data type	Tabular data	Images and sound
Amount of data	From several dozen to several hundred records	At least one thousand to several tens of thousands of records
Explainability	XAI (explainable AI)	Black box
Output	Classifications, forecasting, and action plans	Classifications and forecasting
Hardware requirements	Computer with a generic CPU (Even a notebook can be used)	Parallel computers such as GPUs

Wide Learning has been applied to industrial fields such as manufacturing, finance, and healthcare. However, a gap separating explainable AI and humans apparently remains: AI's explanations and humans' understanding. Sometimes the explanations which Wide Learning offers are difficult for humans to understand. Therefore, some experts must reinterpret them or summarize them for non-experts to understand them easily, which requires great expenses in terms of costs and time in field operations.

Project overview

In the project you will develop an enhanced view of explanations of Wide Learning assisted by mathematics. For purposes of enhancement, you should consider several aspects such as those below.

1. Convincingness: Explanations must be convincing and accepted by users.

Therefore, evaluating or estimating the explanations' likelihood of convincing users is important. That evaluation will require mathematical modeling of the explanation's characteristic of "convincingness". The simplest modeling would be to use the length of an explanation based on the assumption that a short explanation is easy to understand. However, better modeling might be devised by assuming more practical recognition models.

2. **Variety:** An explanation for a judgment is not always unique. Often, more than one perspective is necessary to explain a judgment. The phenomenon that many equally good models explain some given data well is called the Rashomon effect [5]. A set of many equally good models is called a Rashomon set [6]. Although a Rashomon set gives variety in explanation, a tradeoff between variety and simplicity (convincingness) should be considered.
3. **Discoverability:** When an explanation is not convincing to a user, it is not always because it is a bad explanation. It might be a new finding that a user has never noticed. In that sense, an explanation must also be evaluated in terms of discoverability.

By exploring one or more aspects of enhancement, you will develop a new way of giving explanations of Wide Learning.

Expectations

When participating in this project, to enhance the explainability of Wide Learning, you will be expected to devise a new user interface of Wide Learning and then implement it. The implementation should be scalable to a large number of data items, for example, more than 100 data items. We look forward to welcoming students who are interested in the Wide Learning Website "Hello, Wide Learning!" [2]. In addition, students who are interested in answers to the following questions are welcomed.

- ✓ What is explainable AI?
- ✓ What is the gap separating AI and humans?
- ✓ What are good interactions between AI and humans?
- ✓ How can AI help humans discover new findings?

Requirements

Programming skills in python and, preferably, C or C++.

References

- [1] S. R. Islam et al., "Explainable Artificial Intelligence Approaches: A Survey," arXiv:2101.09429, 2021.
<https://arxiv.org/abs/2101.09429>
- [2] Wide Learning Website, "Hello, Wide Learning!".
<https://widelearning.labs.fujitsu.com/en/>
- [3] H. Iwashita et al., "Efficient Constrained Pattern Mining Using Dynamic Item Ordering for Explainable Classification," arXiv:2004.08015, 2020.
<https://arxiv.org/abs/2004.08015>
- [4] A. Krizhevsky et al., "ImageNet Classification with Deep Convolutional Neural Networks," In Advances in Neural Information Processing Systems, 25, 2012.
- [5] L. Breiman, "Statistical modeling: The two cultures (with comments and a rejoinder by the author)," Statistical Science, 16(3): 199-231, 2001.
- [6] R. Xin et al., "Exploring the Whole Rashomon Set of Sparse Decision Trees," In Proceedings of the 36th Intl. Conf. on Neural Information Processing Systems, 2022.