Explainable AI Through Combination of Deep Tensor and Knowledge Graph

● Masaru Fuji ● Hajime Morita ● Keisuke Goto ● Koji Maruhashi
● Hirokazu Anai ● Nobuyuki Igata

One of the most significant advancements made in AI in recent years is the greatly enhanced accuracy of machine learning through deep learning. However, because deep learning deals with huge volumes of data and involves vast neural networks in the learning process, it is often difficult to explain how or why an output was reached even if the inference was correct. This issue impedes the application of AI technology to such business areas as finance and medicine, which demand absolute reliability. To address this issue, we have developed an AI technology that combines Deep Tensor, Fujitsu's original machine-learning technology based on enhanced deep learning, and another Fujitsu-developed machine-learning technology based on knowledge graph, a knowledge base represented by graph data taken from past documents and databases. This has enabled us to logically explain the reasons and bases in which Deep Tensor reaches its inference output. This paper explains the technology that makes Explainable AI possible in terms of application cases in network intrusion detection and genomic medicine.

1. Introduction

Recent years have seen amazing progress in AI with results in some fields actually surpassing the intelligence and actions of humans. At the ImageNet Large Scale Visual Recognition Challenge (ILSVRC), an annual competition where participants compete for the best accuracy in image recognition, accuracy improved significantly by deep learning from 2012 onward, and accuracy by AI at long last surpassed that of humans in 2015. For the board games *shogi* and *go*, which have traditionally been considered too complex for machines to defeat humans, AI triumphed over active professional players in *shogi* in 2013[1)] and *go* in 2016.[2)] These victories demonstrated that AI had advanced to a point that it could generate new moves that humans had not yet discovered.

Furthermore, in addition to improving accuracy, AI technologies and services that are increasingly easier for humans to use and operate in a collaborative manner have begun to appear. For example, chatbots return a reply after having an actual conversation with a user. This technology is clearly different from the conventional method in which a user seeking information had to specify a clear objective or search method beforehand. In short, chatbots show how AI can provide added value by forming a more intimate relationship with humans.

Additionally, in the field of machine learning, an important element of AI, development has begun on technology that does not simply make inferences but also presents reasons or a basis for a certain inference. Conventional machine learning technologies have often been of the "black box" type that prevents humans from understanding how a particular inference was obtained. Such technologies are difficult to apply to critical decision-making involving, for example, human life. However, if the basis for an inference can be presented, this problem can be solved and the application of AI to an even broader range of fields can be expected.

Under the above situations, Fujitsu Laboratories has developed two machine-learning technologies: Deep Tensor[3)] and one based on knowledge graph.

The knowledge graph technology converts diverse types of knowledge data in different fields to data in a uniform graph format that can be used as a single and integrated knowledge base (knowledge graph). Integrating knowledge in this way instead of handling it separately as done in the past has made it possible to describe complex relationships that have so far been difficult to depict.

Deep Tensor is a Fujitsu original machine-learning technology that enhances deep learning by enabling learning on graph-structured data. In the past, it was necessary to have an expert design the features used in graph-data learning, but this approach placed limits on accuracy. Deep Tensor expresses graph data in a mathematical format called a tensor and extracts graph-data features by converting that tensor to a uniform tensor representation using tensor decomposition.[4] It also automates the design of features through optimizing the tensor-decomposition process by using extended error backpropagation, an extension of existing neural-network learning technology (**Figure 1**). This approach enables high-accuracy inference.

We have developed technology for logically explaining the reasons or basis for machine-learning inferences by combining our Deep Tensor and knowledge graph technologies. In this paper, we first describe the black-box problem in conventional deep learning. We then describe two technologies that we developed to resolve this problem. The first technology identifies factors within the input data that significantly contribute to the inference result. We overview this technology and describe its effectiveness through an application example in the field of security. The second technology forms a basis for the inference by associating those factors with knowledge in a knowledge graph. We overview this technology and describe its effectiveness through an application example in the field of genomic medicine.

2. Black-box problem in machine learning

In problem solving, in which an expert collaborates with AI, the expert must understand the inference result from machine learning and make a decision on the basis of that result.

Machine learning based on deep learning can achieve high-performance processing, but learning with a large-scale neural network can produce an extremely complicated relationship between input and output. There is consequently no way to explain to the user why that result was obtained.

Explaining the basis for an inference would normally require an expert to search out evidence for that finding in academic literature or elsewhere. In particular, an expert investigating phenomena in which relationships are only partially known would have to search out and associate evidence in academic literature or databases.

For example, considering the application example in genomic medicine described later, an expert would normally have to investigate medical and medical-treatment literature related to the inference result. However, searching for related evidence in articles has traditionally required considerable labor since more than 100,000 articles are published per month in the fields of medicine.

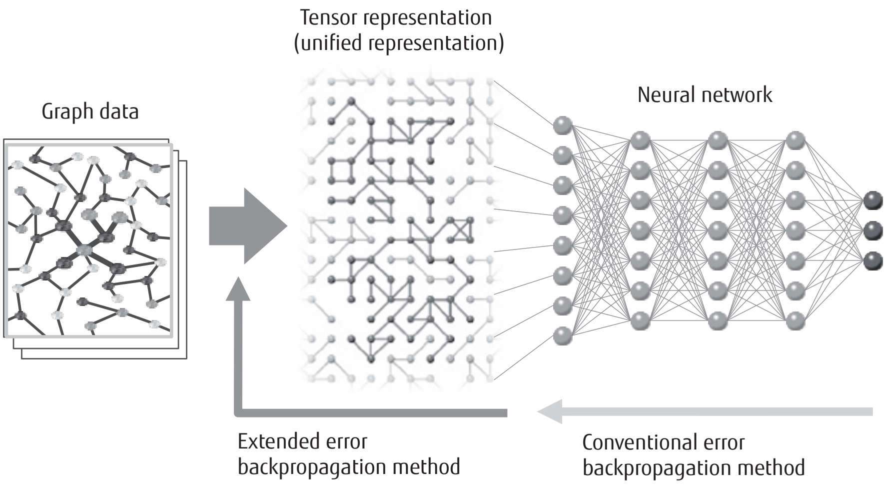


Figure 1
Deep Tensor learning process.

In recent years, research in identifying those portions of input data that significantly contribute to an inference result has been quite active throughout the world. However, in the field of image recognition, this research has only reached the point of simply explaining what parts of an image contribute to the inference result. Moreover, even in cases in which those parts that have a large effect on the inference result can be identified, it is still not possible to present the basis for arriving at that inference.

3. Developed technology

Fujitsu Laboratories developed technology for presenting the reasons or basis for an inference by merging our Deep Tensor and knowledge graph technologies. This technology consists of two main steps. In step 1, it identifies inference factors (partial graphs) that contribute greatly to the inference result in Deep Tensor. In step 2, it establishes a correspondence between those factors and the nodes in a knowledge graph. Furthermore, it connects those factors in the knowledge graph, configures that set of information as the inference basis, and presents the basis to the user. This scheme solves the problem of black-box machine learning while maintaining the high inference accuracy achieved by Deep Tensor.

The following describes these two steps in more detail.

3.1 Identification of inference factors by Deep Tensor

Deep Tensor performs learning by converting graph data into a uniform tensor representation by tensor decomposition and inputting this representation into a neural network. It should be emphasized that the structure of this representation facilitates extraction of those distinguishing factors in the input graph that make a particular contribution to the inference since the representation is learned in such a way as to enhance estimation accuracy. In addition, the fact that this representation is obtained by a linear transformation from the input graph means that the graph itself can easily be obtained from the representation by an inverse linear transformation.

By using this property, we identified the inference factors in the input graph that greatly contributed to the inference result and performed an inverse linear transformation on those factors. As a result, we successfully developed technology for identifying factors that greatly contribute to the inference result from the tensor representation.

In this sense, Deep Tensor can be viewed as a technology that can be applied not only to making high-accuracy inferences but also to identifying inference factors. A variety of applications using Deep Tensor can therefore be expected such as advanced analysis and support of human decision-making.

3.2 Basis forming by using knowledge graph technology

Deep Tensor, which uses graph data for learning and making inferences, has a high affinity with knowledge graph technology for constructing graph data. A knowledge graph consists of a huge amount of graph data that includes all sorts of knowledge. Graph data for learning and inferring can therefore be provided to Deep Tensor by extracting partial graphs from the knowledge graph.

Additionally, by using graph data based on knowledge graph technology and producing an inference and identifying inference factors by Deep Tensor, it becomes possible to read out factors that have a significant influence on the inference result as actual knowledge stored in a knowledge graph.

This basis-forming ability of the knowledge graph technology forms the basis for a knowledge path by making appropriate associations among items of knowledge influencing the inference and connecting the target of inference with the inference result along that path.

Since a knowledge graph includes all sorts of knowledge data, a vast amount of knowledge can be presented by following associations from an item of knowledge on the knowledge graph. When linking the reasons for the inference with the inference result, a knowledge graph forms a basis by tracing those associations. However, a huge number of paths may exist between the two items of knowledge on the knowledge graph. As a consequence, using unrelated information as constituent elements when forming a basis can result in an unsuitable basis for explaining the inference result.

To deal with this problem, the knowledge graph technology searches out a graph structure using inference factors from Deep Tensor as a clue so that only knowledge that is highly relevant to the identified

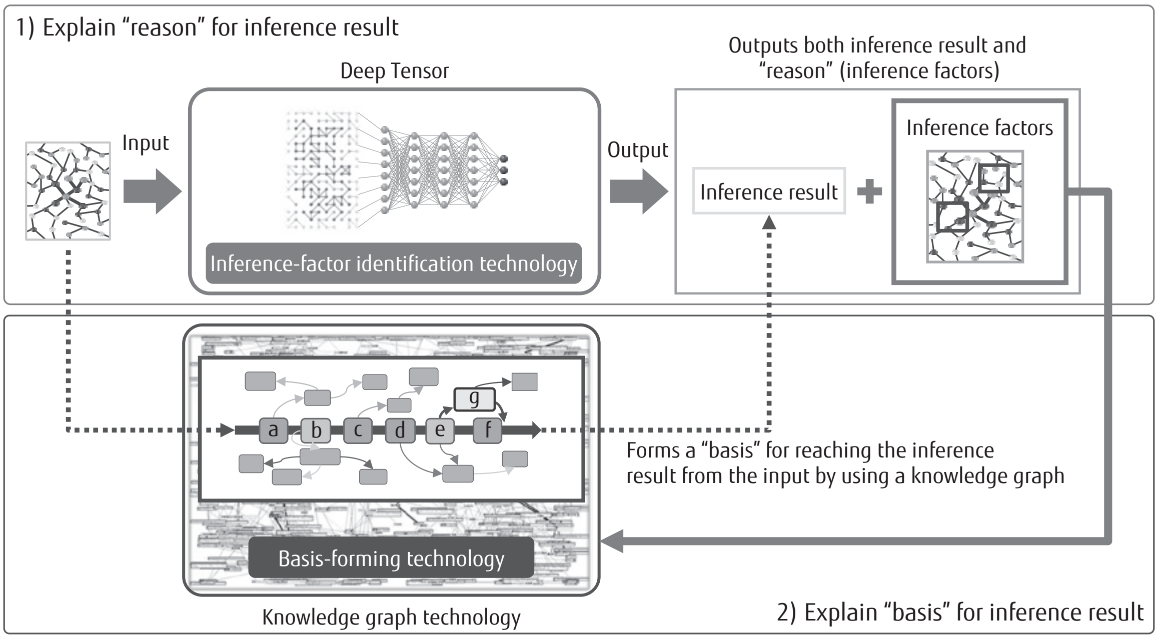


Figure 2
Basis-forming technology combining Deep Tensor and knowledge graph technologies.

inference factors are extracted in forming the basis (**Figure 2**).

4. Evaluation and application examples

We evaluated the practical use of the developed technology. We first evaluated the inference-factor identification technology by applying it to the field of network intrusion detection. We then evaluated the overall Explainable AI that integrates the inference-factor identification and basis-forming technologies by applying it to the field of genomic medicine. In the following subsections, we describe each of these evaluations.

4.1 Evaluation of inference-factor identification technology

We performed this evaluation using the benchmark data sets[5] for network intrusion detection released by the Defense Advanced Research Projects Agency (DARPA) in the United States. These data sets include, as correct factors, log entries that are confirmed to be actual intrusions with respect to network communication log entries detected by an intrusion detection system.

To begin with, we used the inference-factor identification technology, Deep Tensor, to extract the top three factors that significantly contributed to an inference from the data. An example of this process is shown in **Figure 3**. In the figure, the symbol "★" indicates correct factors that are related to the actual intrusion.

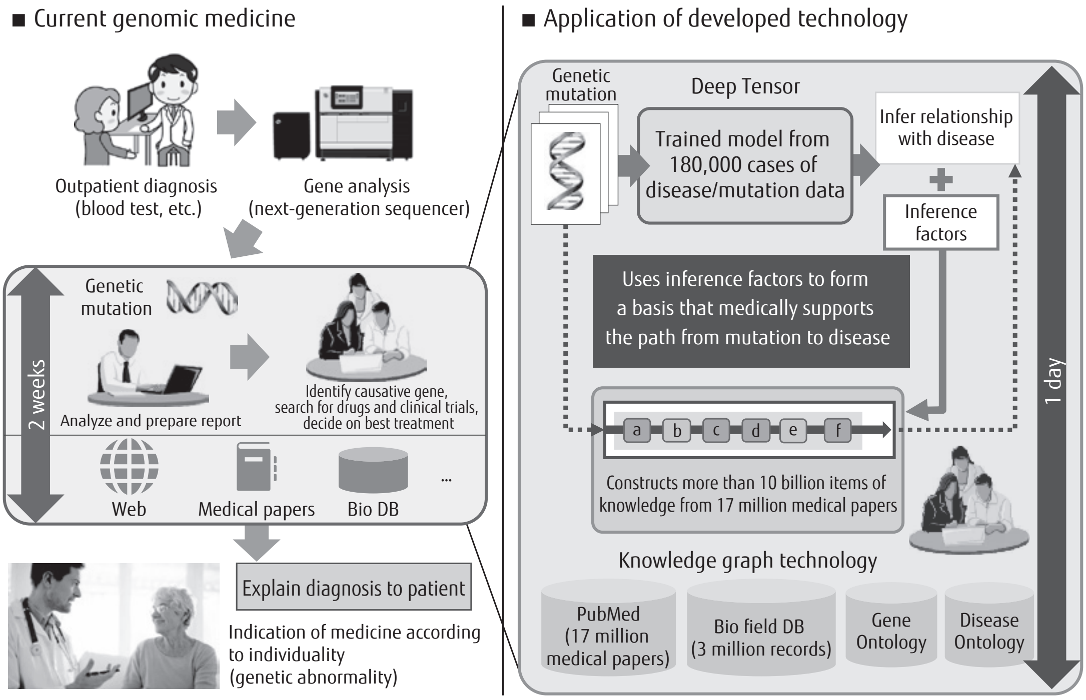
We also evaluated the rate at which extracted inference factors matched correct factors (accuracy) by comparing with that of an existing technology. The accuracy of the existing technique ranged from 10% to 28% while that of Deep Tensor ranged from 23% to 52%. The result indicates that the latter's inference-factor identification technology could obtain an accuracy higher than that of the former. This result suggests that Deep Tensor can be viewed as a machine-learning technology with high explanatory power suitable for identifying inference factors.

4.2 Evaluation of Explainable AI

In order to evaluate Explainable AI, which integrates inference-factor identification technology and basis-forming technology, we performed a simulation experiment in relation to improving the efficiency of survey work by experts in the field of genomic medicine. Using a knowledge graph constructed from public

Jul 29 18:50-19:00 (total of 612 log entries)							
Sending IP address (10 types)		No. of entries	Contribution	Receiving IP address (15 types)		No. of entries	Contribution
★	207.253.084.013	6	0.636	★	172.016.118.020	6	0.636
	172.016.114.207	66	0.207		207.037.252.205	66	0.207
	207.230.054.203	117	0.001		206.132.025.051	27	0.005
Sending port (371 types)		No. of entries	Contribution	Receiving port (246 types)		No. of entries	Contribution
★	49724	6	0.636	★	584	3	0.410
	8739	1	0.003		80 (http)	130	0.186
	8328	1	0.003	★	410	2	0.181

Figure 3
Example of extracting inference factors.



Prepared with Reference 7)

Figure 4
Effect of applying developed technology to genomic medicine.

databases in the field of bioinformatics and a medical-literature database,[6)] we searched for knowledge that could provide corroborating evidence for phenomena in which relationships are only partially known and checked to see whether links could be established (**Figure 4**).

To begin with, we used Deep Tensor to learn relationships between genetic mutations and pathogenicity from public databases. We then extracted information and academic papers related to factors determined by inference-factor identification technology and formed a basis. The basis-forming example of **Figure 5** shows the genetic mutation targeted for inference as a pentagonal-shaped node, factors significantly contributing to the inference result as circle nodes, academically supporting knowledge extracted from medical literature as square nodes, and disease candidates as triangular nodes.

In the figure, the edges (solid lines) connecting nodes indicate that those items of knowledge are

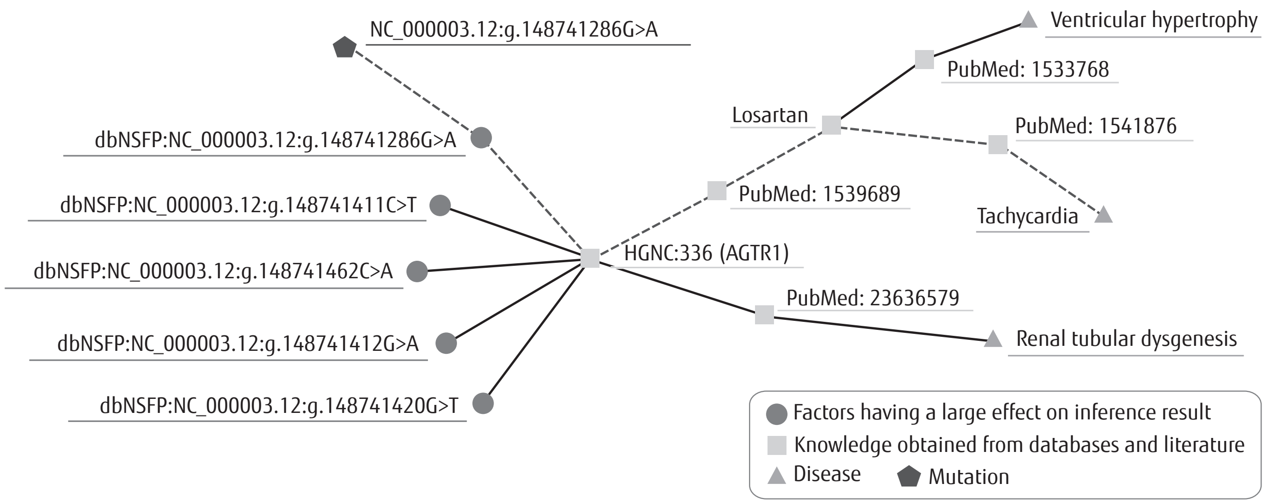


Figure 5
Example of forming a basis.

related on the knowledge graph. The broken lines, on the other hand, interconnect the genetic mutation and its gene, the drug Losartan targeting that gene, and the disease related to that drug thereby presenting the relationship between the genetic mutation and a disease. In short, by starting with the genetic mutation and tracing out the relationship between genes and drugs and that between drugs and diseases on the knowledge graph, it becomes possible to construct a graph of associated knowledge that extends to candidate diseases as a basis that can be viewed by the user.

In the example of Figure 5, the user can judge the relationship between the genetic mutation and the disease Tachycardia by reviewing only the two papers situated on the lines connecting the genetic mutation and that disease without having to survey all 34 papers related to genetic mutation NC_000003.12:g.148741286G>A. This technology was therefore able to greatly reduce the labor involved in identifying and testing for diseases related to a genetic mutation.

5. Conclusion

This paper described AI technology that combines our Deep Tensor and knowledge graph technologies to logically explain the reasons or basis of an inference obtained by Deep Tensor.

A portion of the data used in validating the effectiveness of this technology in the field of genomic medicine was an outcome of joint development with Kyoto University in relation to “Construction of Clinical Genome Knowledge Base to Promote Precision Medicine” in the “Program for an Integrated Database of Clinical and Genome Information” under the Japan Agency for Medical Research and Development (AMED).

Looking to the future, we plan to apply this technology to genomic medicine and other fields to quantitatively evaluate its ability in explaining the basis of an inference and to evaluate its overall usefulness. In the field of genomic medicine, we plan to enlist the cooperation of research institutions related to medical treatment to evaluate whether academic bases presented by this technology are convincing to experts and whether they are sufficiently easy to understand. We also plan to apply the technology to other fields such as finance, where it could be used to check the validity of an inference using knowledge of rules and regulations when teaching an AI system to make automatic loan evaluations.

Finally, we plan to expand our knowledge graph technology in a variety of fields and conduct proof of concept (PoC) demonstrations with the aim of commercializing this technology in FY2019 in the form of services as part of FUJITSU Human Centric AI Zinrai.

References

1) DWANGO, Japan Shogi Association: Second Shogi Den-O-Sen. March 2013.

2) Google: "AlphaGo: Mastering the ancient game of Go with Machine Learning." January 2016.
https://ai.googleblog.com/2016/01/alphago-mastering-ancient-game-of-go.html
3) K. Maruhashi et al.: Learning Multi-way Relations via Tensor Decomposition with Neural Networks. Thirty-Second AAAI Conference on Artificial Intelligence (AAAI-18), February 2018.
4) T. G. Kolda et al.: Tensor decompositions and applications. SIAM Review, Vol. 51, No. 3, pp. 455–500, 2009.
5) MIT Lincoln Library: DARPA Intrusion Detection Data Sets.
http://www.ll.mit.edu/r-d/datasets/
6) US National Library of Medicine: PubMed.
https://www.ncbi.nlm.nih.gov/pubmed/
7) Hokkaido University Hospital: Information on establishment of outpatient genetic testing for cancer (for patients and medical practitioners). (in Japanese).
http://www.huhp.hokudai.ac.jp/hotnews/detail/00001144.html

Masaru Fuji
Fujitsu Laboratories Ltd.
Dr. Fuji is currently engaged in AI-related research such as knowledge graph technology and natural language processing.

Hajime Morita
Fujitsu Laboratories Ltd.
Dr. Morita is currently engaged in research of natural language processing.

Keisuke Goto
Fujitsu Laboratories Ltd.
Dr. Goto is currently engaged in research and development of machine learning in relation to string indexing and graph data.

Koji Maruhashi
Fujitsu Laboratories Ltd.
Dr. Maruhashi is currently engaged in research and development of data mining and machine learning in relation to graph data.

Hirokazu Anai
Fujitsu Laboratories Ltd.
Dr. Anai is currently engaged in research and development of formula manipulation, mathematical optimization, and artificial intelligence.

Nobuyuki Igata
Fujitsu Laboratories Ltd.
Mr. Igata is currently engaged in research of information retrieval, knowledge processing, and natural language processing.