

Fujitsu, Kyoto University Develop Explainable AI Verification System for Estimating Disease-Causing Potential of Genetic Mutations

Tokyo and Kyoto, Japan, March 11, 2021 - Fujitsu Limited and a Kyoto University research group led by Professor Yasushi Okuno of the Kyoto University Graduate School of Medicine, has developed an AI verification system called "MGeND Intelligence." When the genetic mutation (Note 1) information of a patient is input into this system, its ability to cause disease is estimated by pathogenicity estimation AI using machine learning technology and the explanatory text of the basis of this finding is generated and displayed together with the estimated result by the explainable AI. This explanation offers a useful reference to doctors who are considering a treatment plan or genomic medical researchers. The system draws from the "Integrated Database of Clinical and Genomic Information program (Note 2), managed by the Japan Agency for Medical Research and Development, and from April 2021, Kyoto University plans to offer MGeND Intelligence to joint researchers and institutions.

In addition to the pathogenicity estimation AI, the MGeND Intelligence verification system includes explainable AI that shows the basis of the system's findings, as well as literature search AI that supports the retrieval of related articles. Working with the Integrated Database of Clinical and Genomic Information, "MGeND," which was made public by Kyoto University in 2018, Fujitsu supports the research and clinical interpretation of genetic mutations, including mutations with unknown pathogenicity, by medical professionals and researchers.

The use of this verification system ultimately offers the potential to deliver greater innovations in medical care for areas including treatment planning for genomic therapies for illnesses like cancer, accelerating the optimization of patient-centric medical care.

Background

In genomic medicine, knowledge about whether a genetic mutation in a patient has the ability to cause disease is essential. If researchers can locate a pathogenic mutation in a patient's gene, they can consider a treatment for that specific mutation. However, only a small percentage of the vast number of possible genetic mutations have been linked to disease. Genetic mutations with unknown pathogenicity present a challenge in that they do not provide useful information for treating disease.

Since November 2016, Fujitsu and Kyoto University have been participating in the Integrated Database of Clinical and Genomic Information Program promoted by the Japan Agency for Medical Research and Development, engaging in research and development to support the work of medical professionals and researchers in examining genetic mutations with AI and machine learning technologies. Fujitsu and Kyoto University have now developed an AI system that can estimate the presence or absence of disease causing potential for unknown genetic mutations and explain the basis for this finding. Kyoto University will make MGeND Intelligence available to joint researchers and collaborating institutions.

Characteristics of the MGeND Intelligence verification system

"MGeND Intelligence" is a verification system jointly developed by Fujitsu and Kyoto University, designed to support researchers studying genetic mutations. This was achieved through the application of three technologies possessed by Fujitsu Laboratories, Ltd.: machine learning technology trained on graph structure data (Note 3) that can even express complex events; explainable AI technology that can clarify the basis of findings; and natural language processing technology that can accurately extract relationships such as mutations and diseases from a large corpus of textual data.

When the genetic mutation information of a patient is input into the system, its ability to cause disease is estimated by the pathogenicity estimation AI using machine learning technology, and the explanatory text of the basis of this finding is generated and displayed together with the estimated result by the explainable AI. In addition, medical professionals and researchers can gather the basis of their estimation results by themselves. Information in public databases related to mutations and the certainty of estimation by AI can be displayed by multi-faceted visualization together with information on surrounding mutations, while researchers can use the literature search support AI, which leverages natural language processing technology, to search for articles related to target mutations and identify and display the descriptions.

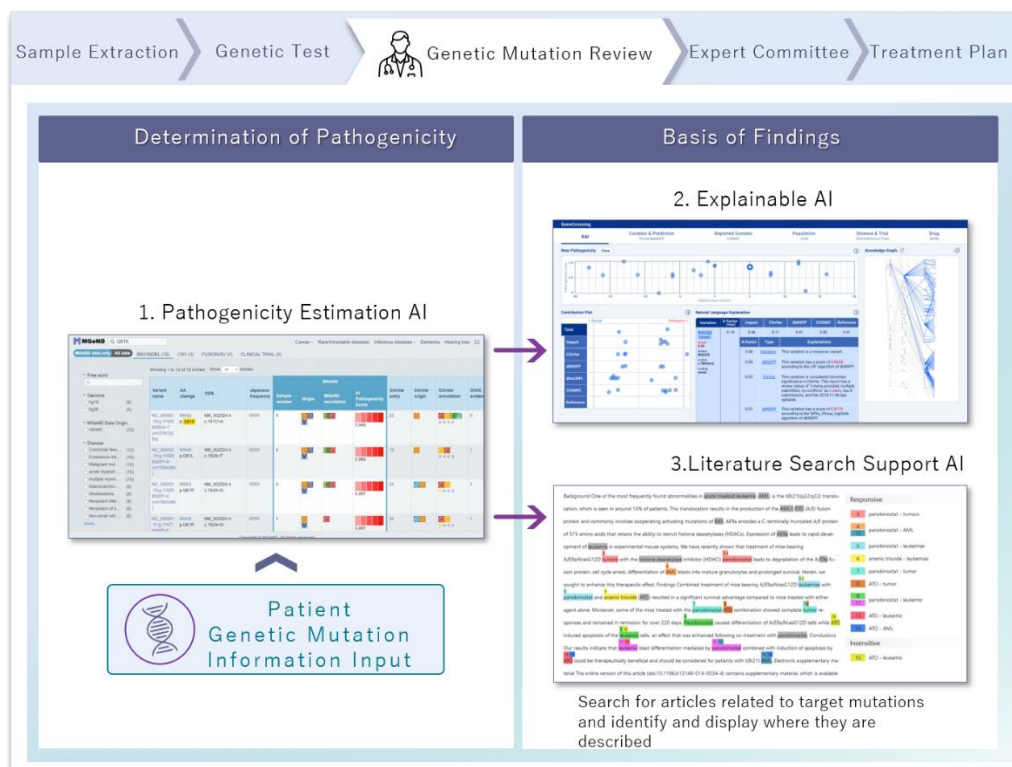


Fig. 1 System Overview

1. Highly Accurate Estimation of Pathogenicity of Genetic Mutations

Based on information from public databases related to genetic mutations whose pathogenicity is known, a Gene Mutation Knowledge Graph (Note 4) was created--this serves as machine learning training data. In constructing the Knowledge Graph, a group led by Professor Yasushi Okuno of Kyoto University defined an effective graph range for estimation, and using Fujitsu Laboratories' "Deep Tensor" (Note 5) machine learning technology for training graph structure data, an AI was created that can accurately estimate the presence or absence of pathogenicity for gene mutations.

2. Explain in natural text the rationale for the AI's finding in a way easy for medical professionals and researchers to understand

With estimation using "Deep Tensor," it is possible to show the graph node that greatly contributed to the finding on the knowledge graph (Note 6). However, since the determination of pathogenicity is carried out comprehensively based on multiple factors, it was not enough to explain the pathogenicity estimation simply by showing the part which became the basis of the finding on the knowledge graph. Therefore, interviews were conducted with Professor Okuno's group and genomic medicine researchers to determine what kind of explanation should be used to reliably estimate pathogenicity using AI and extracted multiple explanation patterns. Then, the degree of fit

of each explanation pattern is calculated from the estimation factor obtained through the "Deep Tensor," and a sentence explaining the basis of the finding is generated based on multiple patterns with a high degree of fit. Medical professionals and researchers can understand the important evidence that has contributed significantly to the estimation just by reading the explanation of the estimation.

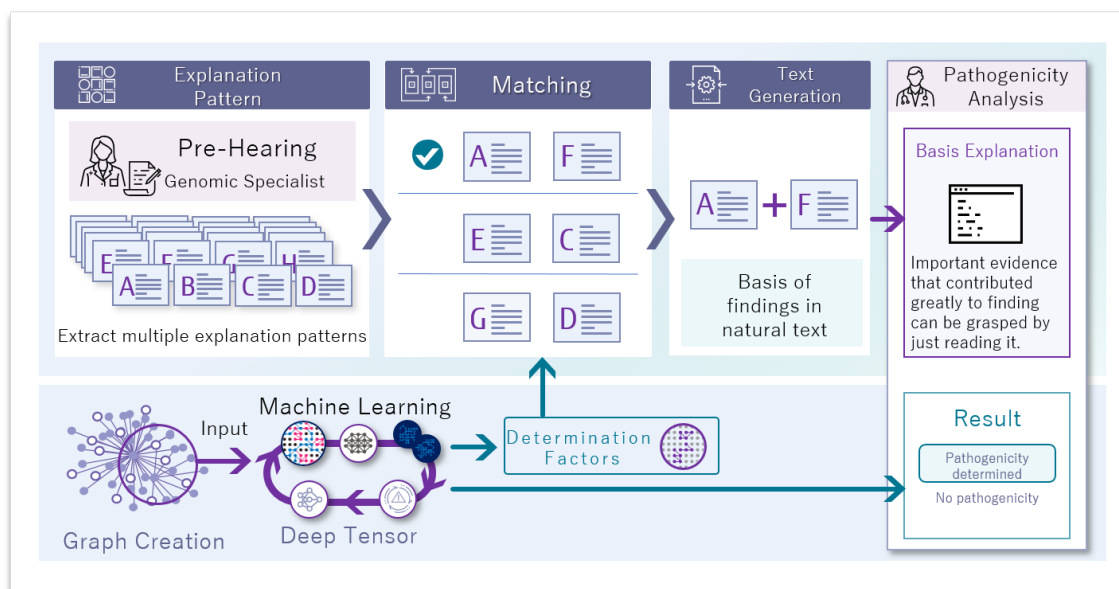


Fig. 2 Flow chart of generation of explanatory text of finding basis

3. Helping Investigate Underlying Articles to Determine Pathogenicity

In a large number of articles, the part describing the relationship between the target mutation and an arbitrary disease is searched, and the relevant part in the article is highlighted. For related information, for example, a large body of literature can explicitly display only the text that describes the relationship between the mutation and the drug. This makes it easy to examine the basis of findings contained in past papers and greatly reduces the time required to examine the estimation results based on known facts (Note 7).

With the cooperation of the Human Genome Center of the Institute of Medical Science at the University of Tokyo, an evaluation test of the entire pathogenicity estimation verification system was conducted, from the pathogenicity estimation to the acquisition of evidence, for unknown mutations (Note 8) in the treatment of genetic disease, and made suggestions from a clinical perspective reflected in the system. As a result, it was confirmed that the results of the estimation of the pathogenicity for each mutation were reasonable, and that the text explaining the reason was appropriate from the viewpoint of specialists.

For example, the pathogenicity of one mutation (Note 9) was unknown at the time of the creation of the verification system used in the evaluation test, and was later judged to be pathogenic. In this case the pathogenicity estimation AI made a correct estimation of the mutation's pathogenicity. In addition, the evaluators were able to obtain an overview of the bases of the AI's findings by reading the AI-generated explanatory text, and to gather sufficient evidence to consider that it was pathogenic.

Future Plans

Based on these results, Fujitsu will further enhance its explanatory functions and develop new functions to support genomic medicine in hospital information systems, mainly based on the use of electronic medical records. Fujitsu will promote the introduction of these functions in hospitals and medical institutions throughout Japan.

Kyoto University plans to provide MGeND Intelligence mainly to joint researchers and institutions participating in the Japan Agency for Medical Research and Development project.

Notes

[1] genetic mutation : Indicates the genetic diversity. For convenience, the term "genetic mutation" is used here to include disease-independent genetic polymorphisms that are present in a given percentage of the mass.

[2] Integrated Database of Clinical and Genomic Information program : Based on the government's interim report by the Council for the Promotion of Genomic Medicine, this project verifies the relationship between genomic information and disease specificities and clinical characteristics for Japanese people, develop a database that comprehensively handles clinical and genetic information that can be used for clinical and research purposes, and integrally promote advanced research and development utilizing the research infrastructure.

[3] graph structure data : Non-linear data structure consisting of nodes and edge used to express information.

[4] knowledge graph : Collection of information expressed in the form of papers and research results, and is used to store knowledge handled by computers in a collection of graph structures that connect related things.

[5] Deep Tensor : Fujitsu's proprietary machine learning technology enables high-precision analysis of graph-structured data that can express connections between people and objects.

[6] graph node that greatly contributed to the finding on the knowledge graph :

"Fujitsu Fuses Deep Tensor with Knowledge Graph to Explain Reason and Basis Behind AI-Generated Findings" <https://www.fujitsu.com/global/about/resources/news/press-releases/2017/0920-02.html>

[7] greatly reduces the time required to examine the estimation results based on known facts :

“Fujitsu Improves Efficiency in Cancer Genomic Medicine in Joint AI Research with the Institute of Medical Science at the University of Tokyo”

<https://www.fujitsu.com/global/about/resources/news/press-releases/2019/1106-01.html>

[8] unknown mutations: A gene mutation whose pathogenicity was unknown at the time of the model construction of the AI estimation installed in the system used for the evaluation, and whose pathogenicity is now known.

[9] pathogenicity of one mutation : Mutation of Gene MYO7A:R1873Q。

About Fujitsu

Fujitsu is the leading Japanese information and communication technology (ICT) company offering a full range of technology products, solutions and services. Approximately 130,000 Fujitsu people support customers in more than 100 countries. We use our experience and the power of ICT to shape the future of society with our customers. Fujitsu Limited (TSE:6702) reported consolidated revenues of 3.9 trillion yen (US\$35 billion) for the fiscal year ended March 31, 2020. For more information, please see www.fujitsu.com.

About Kyoto University

Kyoto University is one of Japan and Asia's premier research institutions, founded in 1897 and responsible for producing numerous Nobel laureates and winners of other prestigious international prizes. A broad curriculum across the arts and sciences at both undergraduate and graduate levels is complemented by numerous research centers, as well as facilities and offices around Japan and the world. For more information please see: <http://www.kyoto-u.ac.jp/en>.

Press Contacts

Public and Investor Relations Division

Inquiries(<https://www.fujitsu.com/global/about/resources/news/presscontacts/form/index.html>)

Company: Fujitsu Limited