

AI learns complex gene-disease patterns

2 September 2019



An artificial intelligence application is being used by KAUST researcher to find genes that cause diseases. Credit: Kiyoshi Takahase Segundo / Alamy Stock Photo

Artificial intelligence (AI) is being harnessed by researchers to track down genes that cause disease. A KAUST team is taking a creative, combined deep learning approach that uses data from multiple sources to teach algorithms how to find patterns between genes and diseases.

Machine learning uses algorithms and statistical models to identify patterns and associations among data to solve specific problems. By inputting enough known data, like tagged images of "Jack," the system can eventually learn to suggest other nontagged images that include Jack.

Researchers are using this application of AI to find [genes](#) that cause diseases. However, only a limited number of genes have been experimentally confirmed to be causative. This means that scientists do not have a lot of data to input into their programs to help them learn the patterns depicting gene-disease associations. Thus, they need to be creative to find ways to teach machine learning algorithms to learn and then look for these patterns.

Database and information management specialist Panagiotis Kalnis, computational bioscientist Xin Gao and colleagues have developed a deep

learning [model](#) they say outperforms current state-of-the-art methods.

First, they resorted to known databases to extract information on gene locations and functions and on how and when they turn on and off. This data was used to teach algorithms to find genes that work together. Then, they obtained data on the features of genetic diseases from other databases. This taught the algorithms how to identify diseases with similar manifestations. They combined these datasets with data on the known associations between 12,231 genes and 3,209 diseases.

The KAUST model extracts the patterns learned from how genes network and about the similarities among genetic diseases and transfers them to a [deep learning model](#) called a graph convolutional network. This delivers another set of data that is placed in matrices, such as those used in recommendation systems, to predict gene-[disease](#) association.

The model was able to identify complex, nonlinear associations between genes and diseases, allowing it to go on to predict new associations. "By making use of more information, we achieved better accuracy than the state-of-the-art methods currently in use," says Peng Han, the first author of the study. "But, even though we outperformed other methods in our experiments, it is still not accurate enough to be applied to industry," he adds.

The team next plans on improving their model's accuracy by incorporating more kinds of data. They will also apply the method to solve other types of problems where only limited data is available, such as recommending new locations to visit based on a user's past preferences.

More information: Peng Han et al. GCN-MF, *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining - KDD '19* (2019). [DOI: 10.1145/3292500.3330912](#)

Provided by King Abdullah University of Science
and Technology

APA citation: AI learns complex gene-disease patterns (2019, September 2) retrieved 3 December 2022
from <https://techxplore.com/news/2019-09-ai-complex-gene-disease-patterns.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.