ПРИМЕНЕНИЕ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В ДИАГНОСТИКЕ ГЕНЕТИЧЕСКИХ ЗАБОЛЕВАНИЙ

ARTIFICIAL INTELLIGENCE IN DIAGNOSTICS OF GENETIC DISEASES

Дель Посо Мурило М.Э.

Раджкумар Д.С.Р.
Кандидат медицинских наук

Курский государственный медицинский университет

Del Pozo Murilo M. E.

Rajkumar D.S.R.
Candidate of Medical Sciences

Kursk State Medical University

E-mail: rajkumardsr@kursksmu.net

Резюме

Искусственный интеллект помогает ученым с помощью компьютерного программного обеспечения, новых программ и алгоритмов идентифицировать варианты ДНК в наших геномах и ошибки в написании ДНК человека, которые с наибольшей вероятностью могут вызвать заболевания. Эти прогнозы могут помочь быстрее диагностировать даже редкие заболевания и послужить руководством для создания новых лекарств и успешного лечения.

Ключевые слова: генетические заболевания, искусственный интеллект, нарушение обмена веществ, пропионовая ацидемия, прогнозирование генетических мутаций, генетические нарушения.

Summary

Artificial intelligence is helping scientists using computing software, new programs and algorithms to identify the DNA variants in our genomes and misspellings in human DNA that are most likely to cause disease. These predictions may help to diagnose even uncommon diseases more quickly and provide guidance for the creation of new drugs and a succesful treatment.

Key words: genetic diseases , artificial intelligence, metabolic disorder, propionic acidemia , genetic mutation prediction, genetic disorders.

Библиографическая ссылка на статью
Дель Посо Мурило М.Э., Раджкумар Д.С.Р. Применение искусственного интеллекта в диагностике генетических заболеваний // Innova. - 2024. - Т.10 № 1. - С.13-16.

References to the article
Del Pozo Murilo M. E., Rajkumar D.S.R. Artificial intelligence in diagnostics of genetic diseases // Innova. - 2024. - T.10 No. 1. - P.13-16.

A large number of genetic mutations that cause a disorder directly, including those causing sickle-cell disease and cystic fibrosis, tend to alter the protein that they encode's amino acid sequence.[5]

However, just a few million of these "missense mutations" involving a single letter have been found by researchers. Out of the potential over 70 million in the human genome.[5] It can be challenging to interpret missense mutations discovered by researchers and physicians if they have never seen one before. One problem limiting the use of genomics in healthcare is identifying which protein mutations are likely to cause specific health issues [6].

An innovative algorithm can assess a patient's likelihood of having a hereditary illness that raises cholesterol and can lead to early and occasionally catastrophic cardiac problems [11].

A new technique based on the AlphaFold network solves this problem with accuracy. A lot of information is gathered by healthcare systems about people, ailments, and therapies. They gather defined information and organized data, including lab results and patient profiles. Additionally, they have unstructured data for every patient they have treated, including reports and clinical notes that have been freely written by medical experts [2].

The AI is a modified version of DeepMind's AlphaFold software, which uses the chemical composition of human proteins to predict their three-dimensional structure.In order to determine which

missense mutations are prevalent and hence most likely benign and which are rare and potentially hazardous, AlphaMissense was fed DNA data from humans and closely related primates. Simultaneously, the algorithm learned the "language" of proteins by examining millions of protein sequences and determining the characteristics of a "healthy" protein [14].

The AlphaMissense AI network is a step ahead, according to experts creating tools of a similar nature.It is one of several methods in development that are meant to assist scientists and doctors in "interpreting" the genetic makeup of individuals in order to determine the etiology of a disease.[7] However, before being employed in the clinic, instruments like AlphaMissense which is detailed in a September 19 study in Science will require extensive testing [5].

Researchers have created dozens of different computer techniques that can determine whether a variant is likely to cause disease in order to aid in the interpretation of such "variants of unknown significance." AlphaMissense by Google DeepMind integrates current methods for prediction of mutations, which are increasingly utilizing machine learning [8].

Material and methods. Scientists may scan the complete genomes, or all of an organism's DNA, of mice used as models for human diseases using computer software. Finding the genetic abnormalities causing such diseases is the aim, as is creating new avenues for scientists to better employ genetics to create disease cures [1].

They started by examining enormous data sets that described the physical characteristics, or illness phenotypes, of numerous varieties of mice. They then carried out genome-wide association studies, which pinpoint the relationships between genes and disease-specific traits [1].

The AI software finds genes associated in mice with illness features like diabetes and cataracts. Fang downloaded 29 million documents, which the AI software went through and examined to see whether any of the publications discussed a potential gene related to a certain condition [1]. The artificial intelligence software identifies genes linked to diseases like diabetes and cataracts in mice. After Fang downloaded 29 million documents, the AI program went through and looked over them to see if any of the articles mentioned a possible gene linked to a certain ailment [1].

Additionally, AlphaMissense integrates a kind of neural network known as a protein language model, which was influenced by massive language models and trained on millions of protein sequences rather than words. These have shown to be skilled in both protein structure prediction and protein design [19].

Additionally, the researchers employed AlphaMissense to compile a list of all potential missense mutations in the human genome [9]. In human biology, missense variants mutations in the genetic code that cause a different amino acid to occur in proteinsare essential. These minuscule differences possess the capacity to induce extensive biological disturbances [22]. Genetic mutations known as missense variants alter the structure and functionality of human proteins by introducing an alternative amino acid into the chain [23].

Of these, 32% are thought to be pathogenic and 57% are probably benign [10].

Furthermore, AlphaMissense incorporates a type of neural network called a protein language model, which was trained on millions of protein sequences instead of words and was inspired by large language models. These have demonstrated proficiency in both protein creation and protein structure prediction [18].

To create a list of every possible missense mutation in the human genome, the researchers also used AlphaMissense [15].

In propionic acidemia, a rare metabolic disease that affects one in 20,000 to 500,000 people globally. Higher concentrations of the chemical propionic acid are found in the bodies of patients with propionic acidemia, which can lead to organ damage and recurrent hospital stays.[4] There are situations when a liver transplant is required.Nearly 500 different types of genetic, laboratory, and imaging data were gathered by the researchers [4].

The researchers trained the algorithm to identify which elements of the data are specifically linked to the two types of the condition after collaborating with specialists in propionic acidemia disease to develop a mechanism to divide patients into moderate and severe categories [4].

The researchers fed the program fresh patient data after it had been trained. Determining which data categories were linked to the mild versus severe kind of propionic acidemia was a highly successful algorithmic task. Clinicians can concentrate on identifying severe patients more quickly and getting them the care they need as soon as feasible if they have knowledge about which indicators are most closely connected with the severity of propionic academia [4].

Certain diseases can be caused by a single missense variant or a small number of missense variants, but other complicated disorders, like Type 2 diabetes, might be caused by a complex interplay of multiple genetic alterations [21].

Results and their discussion. Artificial

intelligence (AI) subsidiary called DeepMind has identified 89% of 71 million possible "errors" in DNA, compared to just 0.1% found by human scientists. [24]

These technologies enable scientists and clinicians to determine the genetic causes of complicated disorders by forecasting the toxicity of missense variants.[24]

In as little as five hours, researchers at Stanford University School of Medicine have created a technique for quickly sequencing a patient's whole human genome. The researchers used their breakthrough to diagnose rare genetic illnesses in less than eight hours. Their recently developed "ultra-rapid genome sequencing approach" may result in much quicker better clinical laboratory treatments. [12]

Current technologies can detect minor genomic variants, such as insertions or deletions of a short sequence of DNA letters or changes to a single DNA letter. On the other hand, "structural variants" can also be identified as disease causes by GEM. These alterations are more significant and frequently intricate. Ten to twenty percent of hereditary disorders are thought to be caused by structural variations. [13]

In the field of (bio)medicine, data collection and management volume is constantly growing. All of this information must be quickly and effectively gathered, examined, and described.[25]

Based on information on the impact of other closely related mutations, the model has assigned a "pathogenicity score" of 0 to 1 to each of the 71 million potential missense variants. The higher the score, the more probable a specific mutation is to cause or be linked to disease. [3]

Four resources with millions of predictions for missense variants across the human proteome were made available to the research community by the researchers. Out of the 71 million missense variations in the first sample, 32% and 57% were probably benign and harmful, respectively. [19]

Each missense variant in this case had a single nucleotide changed, changing the amino acid. [20]

Conclusions. Gene-level AlphaMissense pathogenicity predictions were the second resource. The third included 19,233 human proteins and 216 million potential single amino acid changes. For 60,000 alternative transcript isoforms, the last and fourth resources provided predictions for every potential missense variant and amino acid replacement for use in future studies. [17] AlphaMissense predictions have the potential to enhance the diagnostic yield of uncommon genetic illnesses, shed light on the molecular impact of variations on protein function, and help identify pathogenic missense mutations and previously unidentified disease-causing genes. Additionally, using structure prediction models, AlphaMissense will support the advancement of specialized protein variation effect predictors. [17]

By merging prediction techniques with the expanding understanding of genetic disorders, artificial intelligence (AI) has the potential to significantly expedite and simplify genome interpretation.[16]

Determining other biophysical properties and stabilizing protein therapeutics can also benefit from the prediction of protein structures. In contrast, the prediction of disease causing variations has generally been of little benefit from addressing structure alone. [26]

References.

1. https://scopeblog.stanford.edu/2022/06/10/using-ai-to-find-disease-causing-genes/.
2. https://www.gaucherdisease.org/blog/ai-and-rare-disease-diagnosis-national-gaucher-foundation/.
3. https://www.wired.com/story/deepmind-ai-alphamissense-genetics-rare-diseases/.
4. https://www.genome.gov/news/news-release/artificial-intelligence-tools-help-scientists-decode-genomic-disorders-and-communicate-genomic-risks.
5. https://www.scientificamerican.com/article/ai-tool-pinpoints-genetic-mutations-that-cause-disease/.
6. https://www.labiotech.eu/opinion/rare-disease-diagnosis-ai/.
7. https://es.wired.com/articulos/la-nueva-ia-de-deepmind-puede-predecir-enfermedades-geneticas.
8. https://www.technologyreview.com/2023/09/19/1079871/deepmind-alphamissense-ai-pinpoint-causes-genetic-disease/.
9. https://www.axios.com/2023/09/20/google-deepmind-ai-genetic-disease.
10. https://phys.org/news/2023-09-google-ai-tool-danger-genetic.html.
11. https://med.stanford.edu/news/all-news/2019/04/ai-identifies-risk-of-cholesterol-raising-genetic-disease.html.
12. https://www.darkdaily.com/2022/11/14/stanford-medicine-scientists-sequence-patients-whole-genome-in-just-five-hours-using-nanopore-genome-sequencing-ai-and-cloud-computing/.
13. https://attheu.utah.edu/facultystaff/ai-technology-identifies-genetic-causes-of-serious-disease/.
14. https://www.theguardian.com/science/2023/sep/19/google-deepmind-ai-tool-assesses-

dna-mutations-for-harm-potential.
15. https://www.darkdaily.com/2023/11/15/google-deepmind-says-its-new-artificial-intelligence-tool-can-predict-which-genetic-variants-are-likely-to-cause-disease/.
16. https://genomemedicine.biomedcentral.com/articles/10.1186/s13073-021-00965-0.
17. https://www.news-medical.net/news/20230920/AlphaMissense-revolutionizes-genetic-mutation-analysis-for-disease-prediction.aspx.
18. https://www.science.org/doi/10.1126/science.adg7492.
19. https://www.wired.co.uk/article/deepmind-ai-alphamissense-genetics-rare-diseases.
20. https://the-decoder.com/deepminds-alphamissense-predicts-potentially-disease-causing-gene-mutations/.
21. https://www.insideprecisionmedicine.com/topics/precision-medicine/alphamissense-classifies-mutation-pathogenicity/.
22. https://www.warpnews.org/artificial-intelligence/new-ai-predicts-genetic-diseases/.
23. https://pharmaphorum.com/news/deepmind-applies-its-ai-expertise-genetic-diseases.
24. https://www.techopedia.com/how-deepmind-alphafold-ai-unveils-hidden-dna-disease-secrets.
25. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6947640/.
26. https://singularityhub.com/2023/09/21/deepmind-ai-hunts-down-the-dna-mutations-behind-genetic-disease/.