AlphaFold

Published On: 02-08-2022

What's in News?

DeepMind, a company based in London and owned by Google, announced that it had predicted the three-dimensional structures of more than 200 million proteins using AlphaFold

Proteins:

Proteins are commonly referred as the **building blocks of life**, made up on some combination of amino acids.

It is easy to identify a protein by its constituent amino acid, but this is just one-dimensional information.

What is important to understand is how these amino acids come together and "fold" to create a protein structure.

For instance, the Sars-Cov-2 has a protein that folds as a spike.

This shape, therefore, is relevant for biologists because so that they can design antibodies and therapeutics to, say, neutralise this protein (thereby taking away its ability to infect more cells).

This three-dimensional information is mostly gathered using cryo-electron microscopes.

Alpha Fold:

AlphaFold is an **AIbased protein structure prediction tool**.

It is based on a computer system called deep neural network.

Inspired by the human brain, neural networks use a large amount of input data and provides the desired output exactly like how a human brain would.

The real work is done by the black box between the input and the output layers, called the hidden networks.

AlphaFold is fed with protein sequences as input.

When protein sequences enter through one end, the predicted threedimensional structures come out through the other.

Working of AlphaFold:

It uses processes based on **"training, learning, retraining and relearning."**

The first step uses the available structures of 1,70,000 proteins in the **Protein Data Bank (PDB) to train the computer model.**

Then, it uses the results of that training to learn the structural predictions of proteins not in the PDB.

Once that is done, it uses the highaccuracy predictions from the first step to retrain and relearn to gain higher accuracy of the earlier predictions.

By using this method, AlphaFold has now predicted the structures of the entire 214 million unique protein sequences deposited in the Universal Protein Resource (UniProt) database.

This update includes predicted structures for plants, bacteria, animals, and other organisms, opening up many new opportunities for researchers to use AlphaFold to advance their work on important issues, including sustainability, food insecurity, and neglected diseases

Implications of AlphaFold:

Knowing protein structure and function is essential to **understanding human diseases**.

Scientists predict protein structures using xray crystallography, nuclear magnetic resonance spectroscopy, or cryogenic electron microscopy.

These techniques are not just timeconsuming, they often take years and are based mainly on trialanderror methods.

AlphaFold has already helped hundreds of scientists **accelerate their discoveries in vaccine and drug development** since the first public release of the database nearly a year back

Implications for India:

Indian scientific community It needs to quickly take advantage of the AlphaFold database and learn how to use the **structures to design better vaccines and drugs**.

For instance, understanding the accurate structures of COVID-19 virus proteins in days rather than years

India will also need to speed up its implementation of publicprivate partnerships in the sciences

The publicprivate partnership between the European Molecular Biology Laboratory's European Bioinformatics Institute and DeepMind made the 25terabyte AlphaFold dataset accessible to everyone in the scientific community at no cost

Learning from this, India could facilitate **joint collaborations** with the prevalent hardware muscle and data science talent in the private sector and specialists in academic institutions to pave the way for data science innovations.

Limitations of AlphaFold:

AlphaFold on its own has limitations, as its designers fully acknowledge.

For example, it is **not designed to predict how a protein's shape is altered by disease-causing mutations**.

It was also **not originally intended to predict how proteins change shape when they interact with other proteins** — although researchers are making progress on this next-generation challenge.

And it's not yet clear whether AlphaFold's predictions will reliably provide the fine-grained detail necessary for drug discovery, such as the precise shape of the area on a protein to which a small molecule might bind — the kind of information that researchers in drug development crave.