A new AI model can accurately predict human response to novel drug compounds

October 17 2022

An illustration of personalized drug responses. Credit: CODE-AE illustration

The journey between identifying a potential therapeutic compound and Food and Drug Administration approval of a new drug can take well over a decade and cost upward of a billion dollars. A research team at the CUNY Graduate Center has created an artificial intelligence model that could significantly improve the accuracy and reduce the time and cost of the drug development process.

Described in a newly published paper in *Nature Machine Intelligence*, the new model, called CODE-AE, can screen novel drug compounds to accurately predict efficacy in humans. In tests, it was also able to theoretically identify personalized drugs for over 9,000 patients that could better treat their conditions. Researchers expect the technique to significantly accelerate drug discovery and precision medicine.

Accurate and robust prediction of patient-specific responses to a new chemical compound is critical to discover safe and effective therapeutics and select an existing drug for a specific patient. However, it is unethical and infeasible to do early efficacy testing of a drug in humans directly. Cell or tissue models are often used as a surrogate of the human body to evaluate the therapeutic effect of a drug molecule. Unfortunately, the drug effect in a [disease model](#) often does not correlate with the drug efficacy and toxicity in [human patients](#). This knowledge gap is a major factor in the high costs and low productivity rates of drug discovery.

"Our new machine learning model can address the translational challenge from disease models to humans," said Lei Xie, a professor of computer science, biology and biochemistry at the CUNY Graduate Center and Hunter College and the paper's senior author. "CODE-AE uses biology-inspired design and takes advantage of several recent advances in

x

x

machine learning. For example, one of its components uses similar techniques in Deepfake image generation."

The new model can provide a workaround to the problem of having sufficient patient data to train a generalized [machine learning](#) model, said You Wu, a CUNY Graduate Center Ph.D. student and co-author of the paper. "Although many methods have been developed to utilize cell-line screens for predicting clinical responses, their performances are unreliable due to data incongruity and discrepancies," Wu said. "CODE-AE can extract intrinsic biological signals masked by noise and confounding factors and effectively alleviated the data-discrepancy problem."

As a result, CODE-AE significantly improves accuracy and robustness over state-of-the-art methods in predicting patient-specific drug responses purely from cell-line compound screens.

The research team's next challenge in advancing the technology's use in [drug discovery](#) is developing a way for CODE-AE to reliably predict the effect of a new drug's concentration and metabolization in human bodies. The researchers also noted that the AI model could potentially be tweaked to accurately predict human side effects to drugs.

machine learning. For example, one of its components uses similar techniques in Deepfake image generation."

The new model can provide a workaround to the problem of having sufficient patient data to train a generalized [machine learning](#) model, said You Wu, a CUNY Graduate Center Ph.D. student and co-author of the paper. "Although many methods have been developed to utilize cell-line screens for predicting clinical responses, their performances are unreliable due to data incongruity and discrepancies," Wu said. "CODE-AE can extract intrinsic biological signals masked by noise and confounding factors and effectively alleviated the data-discrepancy problem."

As a result, CODE-AE significantly improves accuracy and robustness over state-of-the-art methods in predicting patient-specific drug responses purely from cell-line compound screens.

The research team's next challenge in advancing the technology's use in [drug discovery](#) is developing a way for CODE-AE to reliably predict the effect of a new drug's concentration and metabolization in human bodies. The researchers also noted that the AI model could potentially be tweaked to accurately predict human side effects to drugs.

machine learning. For example, one of its components uses similar techniques in Deepfake image generation."

The new model can provide a workaround to the problem of having sufficient patient data to train a generalized [machine learning](#) model, said You Wu, a CUNY Graduate Center Ph.D. student and co-author of the paper. "Although many methods have been developed to utilize cell-line screens for predicting clinical responses, their performances are unreliable due to data incongruity and discrepancies," Wu said. "CODE-AE can extract intrinsic biological signals masked by noise and confounding factors and effectively alleviated the data-discrepancy problem."

As a result, CODE-AE significantly improves accuracy and robustness over state-of-the-art methods in predicting patient-specific drug responses purely from cell-line compound screens.

The research team's next challenge in advancing the technology's use in [drug discovery](#) is developing a way for CODE-AE to reliably predict the effect of a new drug's concentration and metabolization in human bodies. The researchers also noted that the AI model could potentially be tweaked to accurately predict human side effects to drugs.

More information: A Context-aware Deconfounding Autoencoder for Robust Prediction of Personalized Clinical Drug Response From Cell Line Compound Screening, *Nature Machine Intelligence* (2022). [DOI: 10.1038/s42256-022-00541-0](#)

Provided by Graduate Center, CUNY

Citation: A new AI model can accurately predict human response to novel drug compounds

3/4