Can automated AI drug discovery platforms increase the confidence of pharma in AI techniques (report by Receptor.AI)
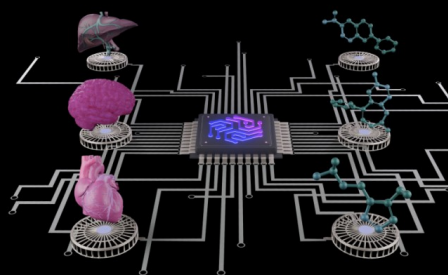
LONDON, UNITED KINGDOM, August 14, 2022 /EINPresswire.com/ -- There is no doubt that Artificial Intelligence is a disruptive technology for drug discovery. The AlphaFold alone, which de facto solved an eternal problem of the protein structure prediction, is enough to claim this. However, there is a certain disappointment in AI drug discovery techniques. The early adopters of AI technologies quickly realised that AI is not a silver bullet which will magically eliminate all the complexity and cost of the drug discovery process.

Receptor.AI has studied and spoken to many parties who, over the last few years, have used AI drug discovery tools and taken their feedback to help overcome these challenges with our solutions.

Early adopters found that the AI drug discovery workflow was weak

The biggest obstacles in AI-based drug discovery are, quite expectedly, the quality of data and poorly determined endpoints. Even such seemingly obvious things as virtual screening become surprisingly complex if

formulated in machine learning terms. We want to classify the compounds into "effective" and "ineffective" for particular diseases, which looks like a trivial ML problem, similar to finding the images of cats among other pictures in computer vision.

However, unlike the images, which either or do not contain the cat's face, we can't even formalise what is "effective" when we speak about molecules. The ultimate goal is to find a compound, which cures the disease, but it's hard physically to get enough training data for this endpoint. That is why R&D teams are forced to use proxy endpoints, such as the strength of binding to a particular target protein.

In a way, all modern AI-based drug discovery searches for the ligands (something that binds) but not for the drugs (something that cures). The binding free energy (or the binding score, which is an approximation of the binding free energy) is a perfect metric for finding the ligands, but not necessarily for finding successful drugs. Indeed, the perfectly binding ligand could fail miserably as a drug due to toxicity, poor ADME properties, harmful off-target interactions and a dozen other non-trivial reasons. That is why Receptor.AI applies AI-based virtual screening and utilises more than 40 predictive models, which estimate the multitude of biological factors, ranging from basic ADME-Tox/PK properties to the proteome-wide off-target interactions assessment.

All these filters are AI models themselves, which require reliable data. A lot of high-quality data! Unfortunately, not all of the data accumulated by big pharma is directly usable for training the AI models. Most of this data is generated by simple high-throughput assays, which are able to perform gargantuan amounts of parallel measurements. However, these measurements usually have low predictive power in terms of the final in vivo endpoint. The AI models trained on these data demonstrate mediocre results not because the model architecture is bad (it is usually state-of-the-art) or the training dataset is too small (it is actually huge), but just because wrong proxy endpoints are used.

Automation and customisation are the key

It is also becoming more and more obvious that there is no such thing as a "universal drug discovery workflow", which is usually referred to as a "drug discovery pipeline" and visualised as a linear sequence of steps.
Based on discussion with multiple biotech and pharma companies, Receptor.AI came to the necessity of automation and fine control over each stage of the AI drug discovery workflow:
• Data preparation and quality assessment of the training and test datasets.
• Correlation of proxy values with biological endpoints and selection of the most significant proxy representation.
• Smart AI assistant of data and model quality control.
• Mapping from the raw data to the most meaningful proxy values.
• Choosing the most suitable AI model architecture and training protocol.
• Performing model training, assessment and tuning.

- Managing different model versions and parameter sets.
- Deploying the model and monitoring its performance in real-world projects.
- Customisable AI pipeline to flexibly approach specific drug discovery projects.

Better tools for overcoming frustration

To address most of these issues, Receptor.AI currently develops the next-generation AI platform for drug discovery, which is not only AI-powered but also AI-automated and AI-assisted.

The platform is fully configurable and includes an easy-to-use pipeline constructor, which allows the user to design its own drug discovery workflow. The platform automates and controls all the routine drug discovery ML tasks focusing on successful drug design, not ligand design: data quality assessment, data filtering, properties and features extraction, model architecture selection, continuous model re-training/tuning triggered by the changes of data or model architecture, deployment of the models, testing their performance, version control of model architectures and parameters, etc.

Such a system solves the majority of technical issues which are repetitively faced by the AI department of the pharma companies in each new project, leading to a clean, effective and organised working environment. Most of the routine tasks are not only automated but are going to be monitored by "advisory AI", which adapts to the internal workflow of the company and balances the tasks and resources accordingly.

What is also important, the platform could be deployed in two operation modes:

1. Completely on-premise, to ensure data protection, access rights and security for big pharma;

2. Cloud-based, which provides almost infinite scalability and accessibility for SMEs.
For now, all major elements of the platform are fully functional in-house and are used in the pilot projects, which Receptor.AI performs in collaboration with several research institutions and CROs. The full system could be deployed on-premise or in the secure cloud and adapted to the needs of the pharma company.

There is also a SaaS solution for super-fast multi-billion scale virtual screening (within an hour), which exposes the most popular modules of the platform by means of a user-friendly interface designed to be "a Google search for novel hit compounds". The SaaS is the most useful for small and medium biotech companies and academic institutions, which require cheap and hassle-free solutions for individual stages of drug discovery workflow.

Alan Nafiiev
Receptor.AI
+ +49 15172837276

email us here
Visit us on social media:
LinkedIn

This press release can be viewed online at: https://www.einpresswire.com/article/586008471