Customer Story: Milner Therapeutics Institute

The data scientists of the Han group at Milner Therapeutics Institute use bespoke machine learning methods to identify new signatures of disease and therapeutic targets, as well as network analysis to gain a deep understanding of the underlying causes of disease. Through research collaborations with academics and industry, their work included studies of cancer, Alzheimer’s, as well as infectious diseases including COVID-19.

This work requires extremely detailed, high-quality data to train their machine learning algorithms.

Without this level of accuracy within their data, their results may not be valid or worth pursuing further. Throughout their research process, they consider where the study might lack data of sufficient depth and quality.
During their COVID-19 research, the team worked to identify drugs that could potentially be repurposed to prevent or treat the infection. The focus was on identifying key proteins and biological pathways that had the potential to be targeted by drugs.

From this research, they identified 200 drugs for repurposing, however, they needed a way to assess risk, implications, and efficacy for each of these drugs.

The team lead Namshik Han was clear about what they were looking for.

“We are very careful about preparing our training datasets. We don’t want to have large quantities of data that are of low quality. For reliable data science, we need the highest quality data.”

Further, they also wanted to be able to look beyond the data to understand its origin, lineage, and updates.

DrugBank was able to provide numerous drug datasets that proved to be powerful for Milner’s research, including targets, toxicity, drug-drug interactions, indications, mechanisms of action, and clinical trials data.

Specifically looking at their work on COVID-19, Milner relied on our drug-drug interaction dataset to understand the implications of each potential drug. Additionally, they integrated the clinical trials dataset as a means of seeing where each of the potential 200 drugs were in the testing process and to make trial information easier to search.
Overall, the integration of DrugBank data into this research has increased the confidence of the Milner team in their research findings. This has resulted in their ability to more quickly and accurately test their hypotheses, validate experiments, and identify potential drugs for combinational therapy.

For their COVID-19 research, DrugBank datasets made it possible for Milner to identify 40 of the 200 drugs predicted to target SARS-CoV-2 as already in clinical trials. Further, they were able to use artificial neural network analysis to classify these 200 drugs into nine distinct pathways, within two mechanisms of action. From here, they were able to narrow their findings further to two drugs that have the potential for preventing infection or treating COVID-19.

“After benchmarking multiple different data sources, we ultimately chose DrugBank because of its high-quality, reliable datasets.”

— Woochang Hwang, Research Associate

Discover more
Request a data sample today!
info@drugbank.com | drugbank.com