Today’s pharmacological research and development uses computing to screen for possible molecules that could become useful constituents in marketable drugs. The industry has traditionally used classical machine learning models based on K-nearest Neighbor clustering, support vector machine (SVM), and other methods for high throughput screening (HTS) followed by wet lab tests to filter for molecules with specific properties.

To date, pharma has explored an incredibly small percentage of the potentially synthesizable drug-like molecules that we know exist today. Within the explored space, a lot has been learned, but discovery has been a slow process using traditional screening methods.

The average pre-clinical drug discovery process takes four years to identify up to 150,000 possible molecules with desirable properties, wet lab screen them, and then lead optimize them down to a set of ten to 20 molecules. The capital investment for this initial process can run up to 2.6 billion dollars before the drug is ready for FDA clearance.¹ The use of AI-directed drug discovery is of significant interest to this industry to help address the incredible amount of time and resource needed throughout the drug discovery process.

**Minds.ai Netrin Molecular Property Predictor Accelerates Discovery**

Many molecules, such as Netrin proteins are of particular interest to pharma. The ability to better map these represents great potential to understand the development and progression of certain diseases and the drugs that may treat them.

Minds.ai’s Netrin Molecular Property Predictor (MPP) is an AI-based tool for exposing the structure-property relationships of molecules, including Netrins. The MPP uses graph convolutional deep neural networks (GCNNs) to make predictions about interactions between molecules and target proteins.

The vision for Netrin is to accelerate the drug discovery process by predicting the outcomes of wet lab tests in seconds compared to the years needed to perform the actual tests. This allows for exploration of millions of more molecules than would practically be possible to perform in the lab. It also allows a parallel approach for multi-variate optimizations, which is not feasible with traditional workflows.
The neural networks are trained on public datasets created from decades of wet lab results and can also be supplemented with private datasets to increase the accuracy of predictions for a particular case.

**Accelerating Netrin Training Performance Up To 2.1X with Intel Technologies**

Training graph convolutional neural networks for molecular models is a CPU- and memory-intensive process. Intel® processors include technologies, such as Intel Advanced Vector Extensions 512 (Intel AVX512) and large memory capacity, that the processors can access directly to accelerate training.

The Netrin MPP algorithms are programmed in Python and use the TensorFlow framework. Minds.ai developers benchmarking training of the Netrin MPP on 2nd Generation Intel Xeon® Scalable processors with and without their algorithms optimized for Intel architecture. Minds.ai engineers used Intel Distribution for Python and applied Intel Optimization for TensorFlow to their algorithms. These enhancements to the TensorFlow framework and Python language leverage the technologies built into the 2nd Gen Intel Xeon Scalable processor architecture and enabled by libraries in the Intel software. The Intel libraries include the Intel Math Kernel Library (Intel MKL) and Intel MKL for Deep Neural Networks (Intel MKL-DNN), all designed to accelerate deep learning on Intel architecture. Running Netrin with these optimizations required no source code changes to the MPP software.

Running on a 2nd Generation Intel Xeon Scalable processor with optimized algorithms accelerated training by up to 2.1X as shown in testing below (Figure 1).

**Graph Convolution Training Performance Using Intel® Optimized Tensorflow on 2nd Gen Intel Scalable Processor**

Figure 1. Training time for optimized Netrin MPP on 2nd Gen Intel Xeon Scalable processor.2

**CURRENT PRE-CLINICAL PROCESS**

- Identify High Throughput Screening (HTS) Targets
- Web Lab Based HTS
- Pharmacokinetics and Safety Screening
- Lead Optimization

<table>
<thead>
<tr>
<th>Stage</th>
<th>Time</th>
<th>Molecules</th>
</tr>
</thead>
<tbody>
<tr>
<td>Identify Targets</td>
<td>~ 1 year</td>
<td>~ 150K molecules</td>
</tr>
<tr>
<td>In-silico HTS, Simulations, Lab Verification</td>
<td>~ 2-3 years</td>
<td>~ 100 molecules</td>
</tr>
<tr>
<td>AI-assisted Lead Optimization</td>
<td>~ 1 year</td>
<td>~ 10-20 molecules sent to trial</td>
</tr>
</tbody>
</table>

- In-silico HTS – via Molecular Property Predictor (MPP) and simulations is dramatically cheaper and faster than web lab HTS
- AI assisted lead optimization will produce unique higher quality molecules to submit to clinical trial.

---

2. In Figure 1, the training throughput is normalized to show the acceleration achieved with Intel Optimization for TensorFlow compared to the baseline TensorFlow training.
Conclusion

Minds.ai Netrin MPP runs up to 2.1X faster on 2nd Gen Intel Xeon Scalable processors with optimizations available in the Intel AI software stack. The Minds.ai solution helps the industry predict molecular properties faster using computer-based deep learning models. Accelerated predictions can reduce the need to run expensive wet lab HTS tests. Netrin MPP can benefit the pharmaceutical industry with:

- Reduced costs
- Faster drug design and development time
- Enables multivariate optimization, as opposed to the traditional sequential and iterative process

Learn More

For more information about Minds.ai, visit https://www.minds.ai

To learn more about the Intel® AI Builders program, visit https://builders.intel.com/ai

Minds.ai is a member of the Intel® AI Builders Program, an ecosystem of industry-leading independent software vendors (ISVs), system integrators (SIs), original equipment manufacturers (OEMs), and enterprise end users, which have a shared mission to accelerate the adoption of artificial intelligence across Intel® platforms.

Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations, and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of the product when combined with other products. For more complete information, visit http://www.intel.com/benchmarks.

Performance results are based on testing as of dates shown in configuration and may not reflect all publicly available security updates. No product can be absolutely secure. See configuration disclosure for details.

Optimization Notice: Intel’s compilers may or may not optimize to the same degree for non-Intel microprocessors for optimizations that are not unique to Intel microprocessors. These optimizations include SSE2, SSE3, and SSSE3 instruction sets and other optimizations. Intel does not guarantee the availability, functionality, or effectiveness of any optimization on microprocessors not manufactured by Intel. Microprocessor-dependent optimizations in this product are intended for use with Intel microprocessors. Certain optimizations not specific to Intel microarchitecture are reserved for Intel microprocessors. Please refer to the applicable product User and Reference Guides for more information regarding the specific instruction sets covered by this notice. Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors.

Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products. For more complete information visit: http://www.intel.com/performance

¹ https://www.policymed.com/2014/12/a-tough-road-cost-to-develop-one-new-drug-is-26-billion-approval-rate-for-drugs-entering-clinical-de.html

² Configuration details:
Testing done September 2019 by Minds.ai.
Platform: Intel® Server Board S26008PB; dual-socket, single-node server
CPU: Intel® Xeon® 6248 CPU @ 2.50 GHz; 20 cores/40 threads; HTT turned on; Turbo boost turned on.
BIOS: SE5C620.868.02.01.0008.031920191559
Memory: 376GB; 6 slots 2666 MTs DDR4 DIMM
Storage: Intel® SSDSC2KB48 480 GB (boot); Intel® SSDSC2KB48 480 GB (application)
Network: Intel® Ethernet Controller X550T 10G
OS: Ubuntu 18.04.2 LTS; Kernel 4.15.0-52-generic
Mitigation variants: Full mitigation
© Intel Corporation. Intel, the Intel logo, and other Intel marks are trademarks of Intel Corporation or its subsidiaries. Other names and brands may be claimed as the property of others.