

# Researchers employing Simulation and Machine-learning techniques for accelerated discovery of new application specific materials

Finding new materials has traditionally been guided by intuition and trial and error,” said Turab Lookman, a physicist and materials scientist in the Physics of Condensed Matter and Complex Systems group at Los Alamos National Laboratory. “But with increasing chemical complexity, the combination possibilities become too large for trial-and-error approaches to be practical.”

Computational and theoretical materials science is playing an increasingly important role in advancing the search for novel materials and understanding the properties of existing ones. Modern computational hardware and software enable faculty to create “virtual laboratories,” where materials are tested and properties predicted computationally.

Researchers are also employing various machine (or statistical) learning methods to accelerate the discovery of new materials. The materials discovery process can be significantly expedited and simplified if we can learn effectively from available knowledge and data.

The Toyota Research Institute (TRI) is investing \$35 million into the artificial intelligence project that would help in the hunt for new advanced battery materials and fuel cell catalysts. TRI Chief Science Officer Eric Krotkov said: “Toyota recognizes that artificial intelligence is a

vital basic technology that can be leveraged across a range of industries, and we are proud to use it to expand the boundaries of materials science. Accelerating the pace of materials discovery will help lay the groundwork for the future of clean energy and bring us even closer to achieving Toyota's vision of reducing global average new-vehicle CO<sub>2</sub> emissions by 90 percent by 2050."

Researchers at the Center for Nanoscale Materials and the Advanced Photon Source, both U.S. Department of Energy (DOE) Office of Science User Facilities at DOE's Argonne National Laboratory, announced the use of machine learning tools to accurately predict the physical, chemical and mechanical properties of nanomaterials.

## **Computational materials screening and targeted experiments reveal promising nitride semiconductors**

Fumiyasu Oba and colleagues at Tokyo Institute of Technology and Kyoto University have used simulations to identify previously undiscovered semiconductors with promising attributes for optical and electronic applications. They used calculations to screen a set of compounds for potential semiconductor candidates. The study identified 11 previously unreported materials, including the particularly promising compound calcium zinc nitride (CaZn<sub>2</sub>N<sub>2</sub>).

The discovery of new semiconducting materials is a scientifically and technologically important issue; increasingly sophisticated electronic devices, such as smartphones and laptops, are raising demand for semiconductors with wider ranges of properties. Silicon and, to a lesser degree, germanium are the foundations of almost all electronic devices that govern modern life. However, these materials are not suited for optoelectronic applications, such as LEDs for

TV or mobile phone screens.

Here, the materials gallium nitride (GaN) and indium nitride (InN) dominate currently, but the discovery of new nitrides could pave the way to new applications. Nitrides tend to be chemically stable and can be readily made with existing techniques. Nitrogen is also a widely abundant and environmentally friendly element, but, at present, the nitrides used in industry are largely limited to gallium and indium compounds.

## **Calcium zinc nitride for optoelectronic applications**

CaZn<sub>2</sub>N<sub>2</sub> has not been reported previously but was identified by the researchers using their materials discovery computational approach. They were also able to predict the correct synthesis conditions for CaZn<sub>2</sub>N<sub>2</sub>. Synthesis of the material using high-pressure techniques confirmed the hypothesised properties and also revealed red luminescence even at room temperature; thereby validating the study's approach.

The paper also shows that other earth-abundant materials, such as calcium magnesium nitride, can be used to tune the electrical properties of CaZn<sub>2</sub>N<sub>2</sub>, further increasing the eligibility of this material for use in devices.

As Oba and colleagues conclude, "The present study demonstrates accelerated materials discovery via cutting-edge computational screening followed by targeted experiments.

## **Machine learning enables predictive modeling of 2-D materials**

In a study published in The Journal of Physical Chemistry Letters, a team of researchers led by Argonne computational scientist Subramanian Sankaranarayanan described their use of machine learning tools to create the first atomic-level model

that accurately predicts the thermal properties of stanene, a two-dimensional (2-D) material made up of a one-atom-thick sheet of tin.

“Predictive modeling is particularly important for newly discovered materials, to learn what they’re good for, how they respond to different stimuli and also how to effectively grow the material for commercial applications—all before you invest in costly manufacturing,” said Argonne postdoctoral researcher Mathew Cherukara, one of the lead authors of the study.

“We input data obtained from experimental or expensive theory-based calculations, and then ask the machine, ‘Can you give me a model that describes all of these properties?’” said Badri Narayanan, an Argonne postdoctoral researcher and another lead author of the study. “We can also ask questions like, ‘Can we optimize the structure, induce defects or tailor the material to get specific desired properties?’”

Unlike most past models, the machine learning model can capture bond formation and breaking events accurately; this not only yields more reliable predictions of material properties (e.g. thermal conductivity), but also enables researchers to capture chemical reactions accurately and better understand how specific materials can be synthesized.

Another advantage of building models using machine learning is the process is not material-dependent, meaning researchers can look at many different classes of materials and apply machine learning to various other elements and their combinations.

## **Accelerating materials property predictions using machine learning**

The standard approaches adopted thus far involve either expensive and lengthy Edisonian synthesis-testing experimental

cycles, or laborious and time-intensive computations, performed on a case-by-case manner.

“Owing to the staggering compositional and configurational degrees of freedom possible in materials, it is fair to assume that the chemical space of even a restricted subclass of materials (say, involving just two elements) is far from being exhausted, and an enormous number of new materials with useful properties are yet to be discovered. Given this formidable chemical landscape, a fundamental bottleneck to an efficient materials discovery process is the lack of suitable methods to rapidly and accurately predict the properties of a vast array (within a subclass) of new yet-to-be-synthesized materials.”

Machine learning methods may be used to establish a mapping between a suitable representation of a material (i.e., its ‘fingerprint’ or its ‘profile’) and *any* or *all* of its properties using known historic, or intentionally generated, data. Subsequently, once the profile property mapping has been established, the properties of a vast number of new materials within the same subclass may then be directly predicted (and correlations between properties may be unearthed) at negligible computational cost, thereby completely bypassing the conventional laborious approaches towards material property determination alluded to above. In its most simplified form, this scheme is inspired by the intuition that (dis)similar materials will have (dis)similar properties.

Researchers are employing machine (or statistical) learning methods trained on quantum mechanical computations in combination with the notions of chemical similarity for efficient and accurate prediction of a diverse set of properties of material systems. Harnessing such learning paradigms extends recent efforts to systematically explore and mine vast chemical spaces, and can significantly accelerate the discovery of new application-specific materials.

# Machine learning based adaptive design strategy accelerates the discovery of new materials

Researchers recently demonstrated how an informatics-based adaptive design strategy, tightly coupled to experiments, can accelerate the discovery of new materials with targeted properties, according to a recent paper published in Nature Communications.

They developed a framework through which starting with a relatively small data set of well-controlled experiments, uses uncertainties to iteratively guide the subsequent experiments toward finding the material with the desired target. “The goal is to cut in half the time and cost of bringing materials to market,” said Lookman. “What we have demonstrated is a data-driven framework built on the foundations of machine learning and design that can lead to discovering new materials with targeted properties much faster than before.” The work made use of Los Alamos’ high-performance supercomputing resources.

The authors have applied their framework to search a shape-memory alloy with very low thermal hysteresis (or dissipation). The functionalities of SMAs, including shape memory effect and super-elasticity, arise from the reversible martensitic transformation between high-temperature austenite and low-temperature martensite phases.

Heating and cooling across the martensitic transformation temperature results in hysteresis ( $\Delta T$ ) as the transformation temperatures do not coincide, giving rise to fatigue. An obstacle for developing low  $\Delta T$  SMAs is the large search space, because a vast majority of transition metals can be alloyed with  $\text{Ni}_{50}\text{Ti}_{50}$

Even in our constrained pseudo-quaternary composition space, there are  $N=797,504$  potential alloys. Such a vast space is difficult to explore with high-throughput experiments or *ab initio* calculations. However, our design loop is able to discover  $\text{Ti}_{50.0}\text{Ni}_{46.7}\text{Cu}_{0.8}\text{Fe}_{2.3}\text{Pd}_{0.2}$  with a low  $\Delta T$  of 1.84 K in the sixth out of nine iterations of our loop.

Prior knowledge, including data from previous experiments and physical models, and relevant features are used to describe the materials. This information is used within a machine learning framework to make predictions that include error estimates. The results are used by an experimental design tool (for example, Global optimization) that suggests new experiments (synthesis and characterization) performed in this work, with the dual goals of model improvement and materials discovery. The results feed into a database, which provides input for the next iteration of the design loop.

We show how we exercise the loop, the key ingredients of which are as follows:

- A training data set of alloys, each described by features and with a desired property (that is,  $\Delta T$ ) that has been measured;
- An inference model (regressor) that uses the training data to learn the feature–property relationship, with associated uncertainties;
- The trained model is applied to the search space of unexplored alloy compositions (for which the property has not been measured), to predict  $\Delta T$  with associated uncertainties;
- Design or global optimization (selector) that provides the next candidate alloy for experiment by balancing the trade-off between exploitation (choosing the material with the best predicted property) and exploration (using the predicted uncertainties to study regions of search

- space where the model is less accurate); and
- Feedback from experiments allowing the subsequent iterative improvement of the inference model.

Lookman and his colleagues focused on nickel-titanium-based shape-memory alloys, but the strategy can be used for any materials class (polymers, ceramics or nanomaterials) or target properties (e.g., dielectric response, piezoelectric coefficients and band gaps). This becomes important when experiments or calculations are costly and time-consuming.

Although the work focused on the chemical exploration space, it can be readily adapted to optimize processing conditions when there are many “tuning knobs” controlling a figure of merit, as in advanced manufacturing applications. Similarly, it can be generalized to optimize multiple properties, such as, in the case of the nickel-titanium-based alloy, low dissipation as well as a transition temperature several degrees above room temperature.

## **Artificial intelligence aids materials fabrication**

A team of researchers at MIT, the University of Massachusetts at Amherst, and the University of California at Berkeley hope to close that materials-science automation gap, with a new artificial-intelligence system that would pore through research papers to deduce “recipes” for producing particular materials.

“Computational materials scientists have made a lot of progress in the ‘what’ to make—what material to design based on desired properties,” says Elsa Olivetti, the Atlantic Richfield Assistant Professor of Energy Studies in MIT’s



Department of Materials Science and Engineering (DMSE). “But because of that success, the bottleneck has shifted to, ‘Okay, now how do I make it?’”

The researchers envision a database that contains materials recipes extracted from millions of papers. Scientists and engineers could enter the name of a target material and any other criteria—precursor materials, reaction conditions, fabrication processes—and pull up suggested recipes.

As a step toward realizing that vision, Olivetti and her colleagues have developed a machine-learning system that can analyze a research paper, deduce which of its paragraphs contain materials recipes, and classify the words in those paragraphs according to their roles within the recipes: names of target materials, numeric quantities, names of pieces of equipment, operating conditions, descriptive adjectives, and the like.

In a paper appearing in the latest issue of the journal *Chemistry of Materials*, they also demonstrate that a machine-learning system can analyze the extracted data to infer general characteristics of classes of materials—such as the different temperature ranges that their synthesis requires—or particular characteristics of individual materials—such as the different physical forms they will take when their fabrication conditions vary.

Much of Olivetti’s prior research has concentrated on finding more cost-effective and environmentally responsible ways to produce useful materials, and she hopes that a database of materials recipes could abet that project.

“This is landmark work,” says Ram Seshadri, the Fred and Linda R. Wudl Professor of Materials Science at the University of California at Santa Barbara. “The authors have taken on the difficult and ambitious challenge of capturing, through AI methods, strategies employed for the preparation of new

materials. The work demonstrates the power of machine learning, but it would be accurate to say that the eventual judge of success or failure would require convincing practitioners that the utility of such methods can enable them to abandon their more instinctual approaches.

## **Machine-learning tool can analyze failed reactions to guide synthesis of new materials**

A team of researchers from Haverford College has developed a machine-learning tool, that by analyzing past data of failed chemical reactions, can guide the choices for synthesizing new materials. The formation of compounds is not fully understood hence synthesis of new materials has relied primarily on exploratory syntheses.

“Failed reactions contain a vast amount of unreported and unextracted information,” says Alex Norquist, a materials-synthesis researcher at Haverford College in Pennsylvania, who is part of the team that has reported the work in *Nature*. “There are far more failures than successes, but only the successes generally get published.” The lessons that can be learnt from failed reactions remain hidden in the lab notebooks of individual researchers.

“Here we demonstrate an alternative approach that uses machine-learning algorithms trained on reaction data to predict reaction outcomes for the crystallization of templated vanadium selenites.”

“We used information on ‘dark’ reactions—failed or unsuccessful hydrothermal syntheses—collected from archived laboratory notebooks from our laboratory, and added physicochemical property descriptions to the raw notebook information using cheminformatics techniques,” the researchers

wrote in a letter to Nature.

They trained an algorithm on data from almost 4,000 attempts to make the crystals under different reaction conditions (such as temperature, concentration, reactant quantity and acidity). That work included transcribing information on dark, failed reactions from the team's archived lab notebooks into a format that a machine could analyse.

"We used the resulting data to train a machine-learning model to predict reaction success." The machine-learning tool was designed to analyze these data and predict whether a set of reagents would yield a crystalline material when combined with a solvent and heated.

## **Decision tree**

Having generated evidence to suggest the algorithm can best human intuition, the team built a decision tree to guide their own decisions in the future. The tree allows researchers to work through questions such as "is sodium present?" to reach a prediction regarding the likelihood of a reaction resulting in a crystal.

To test the algorithm, the team picked out previously untried combinations of reactants, and tried to guess the best processing conditions for making selenite materials. When asked to recommend reactions within this niche, the reaction conditions suggested by the algorithm generated a crystalline product in 89% of the time, compared to this Researchers with a combined 10 years of experience with the materials were right 78% of the time.

The team has set up a website, called the Dark Reactions Project, to encourage others to share – in a machine-readable format – their own failed attempts to make new crystals. "The planning and development of such tools is essential if we are

to eventually make full use of our 'failed' experiments," adds Richard Cooper, a crystallographer at Oxford University, UK.

## References and Resources also include:

- <http://www.fiercebiotech.com/it/machine-learning-tool-mines-lab-notebooks-for-lessons-from-failed-reactions>
- <http://www.sciencenewsline.com/news/2016050920310015.html>
- <http://www.nature.com/ncomms/2016/160415/ncomms11241/full/ncomms11241.html>
- <http://phys.org/news/2016-06-materials-screening-reveal-nitride-semiconductors.html>
- [http://phys.org/news/2016-10-simulations-graphene-defect-s-assets.html?utm\\_source=nwletter&utm\\_medium=email&utm\\_campaign=weekly-nwletter](http://phys.org/news/2016-10-simulations-graphene-defect-s-assets.html?utm_source=nwletter&utm_medium=email&utm_campaign=weekly-nwletter)
- <http://phys.org/news/2016-12-machine-enables-d-materials.html>
- <https://www.engadget.com/2017/03/30/toyota-research-ai-battery-material-hunt/>
- <https://phys.org/news/2017-11-artificial-intelligence-ai-ds-materials-fabrication.html>