A Texas A&M engineering research team is harnessing the power of machine learning, data science and the domain knowledge of experts to autonomously discover new materials.

The team developed and demonstrated an autonomous and efficient framework capable of optimally exploring a materials design space (the materials design space is an abstraction of the concrete world. It is the space of all the possible materials under study, characterized by fundamental material features).

An autonomous system—or artificial intelligence (AI) agent—is defined as any system capable of building an internal representation, or model, of the problem of interest, and that then uses the model to make decisions and take actions independent of human involvement.

The authors of this interdisciplinary work are Dr. Anjana Talapatra and Dr. Raymundo Arroyave from the Department of Materials Science and Engineering, and Shahin Boluki, Dr. Xiaoning Qian and Dr. Edward Dougherty from the Department Electrical and Computer Engineering.

Their autonomous framework is capable of adaptively picking the best machine learning models to find the optimal material to fit any given criteria. Their research, funded by the National Science Foundation and the Air Force Office of Scientific Research, will reduce the time and cost spent going from lab to market by ensuring the greatest possible efficiency in the search for the right material.

The underlying mathematical theory has many applications, including affecting the field of biomedicine. For example, with their Bayesian learning and experiment design framework, a disease can be modeled to uncover critical risk factors to develop effective therapeutics for specific patients and reduce the cost of human clinical trials.

"Advanced materials are essential to economic security and human well-being, with applications in industries aimed at addressing challenges in clean energy, national security and human welfare, yet it can take 20 or more years to move a material after initial discovery to the market.—Materials Genome Initiative

The team wanted to test the framework exhaustively, so they carried out the demonstration in a closed-loop computational platform, using quantum mechanics to predict properties of MAX-phases, which are promising materials for high-temperature applications, including novel oxidation resistant coatings for jet engine turbine blades. The Texas A&M group is also applying the framework to the discovery of high-temperature shape memory alloys that can be used to build aerospace vehicles with morphing wings, for example.
Significant research on efficient experiment design techniques has been done before. However, this team is the first to use a Bayesian based technique (meaning they take stock of all that is known about a material/material class and leverage that knowledge to find the best material) and employ it in an autonomous fashion, continuously searching not only for the next best computation/experiment to run but also for the best model to represent the acquired data.

"The accelerated exploration of the materials space to identify configurations with optimal properties is an ongoing challenge," said Talapatra, who works as a computational scientist in Arroyave’s Computational Materials laboratory. "Current paradigms are centered around the idea of performing this exploration through high-throughput experimentation and/or computation. Those approaches do not account for the constraints in resources available. We have addressed this problem by framing materials discovery as an optimal experiment design."

The methods presented in this research are flexible and adaptable to different research situations. Significantly, Talapatra and Boluki’s algorithm can work with very little initial data, making it ideal for new materials research.

The algorithm represents a smarter step forward compared to previous work in the field. Other algorithms force one to start with a predefined model, which introduces a constraint into the experiment and can skew the results. "Our algorithm can automatically and autonomously decide which model is the best model out of n models, at any given time, depending on the acquired data," said Talapatra. The autonomous computer program reduces the number of steps and limits the use of limited resources. Since it can start with as few as two experiments as initial data points, the algorithm is ideal for optimizing initial experiments and discerning the best path forward.

It can be used as a one-step tool by experimentalists to simply decide on the next material to explore, or as a purely computational tool to replace expensive computational models and reduce computational costs. It can also be used in a combined experimental and computational setup. At the very least, this framework provides a very efficient means of building the initial data set since it may be used to guide experiments or calculations by focusing on gathering data in those sections of the materials design space which will result in the most efficient path to achieving the optimal material.

"Typically, materials research occurs in a very ad-hoc way and serendipity tends to be the rule, rather than the exception," said Talapatra. "The problem is you often don't know the fundamental physics behind why a material is or is not working. Our models are not precise enough. When you start a materials discovery journey, you start with the very basic physical knowledge, such as the number of electrons and what happens when the elements join together. You have to find the similarities between the features and the properties."

"We included as much science as possible in the (artificial intelligence) models," said Boluki, a doctoral student who will defend his thesis next fall. Boluki and Talapatra worked as implementers in the project and coded it in python together.

The paper on the algorithm has been peer reviewed, presented at several conferences and given good feedback from the materials science and engineering community. Engineers and scientists at Texas A&M are already using the program.

**From cell pathology to materials science: The mathematical underpinning**

In 2011, Qian and Dougherty began collaborating on enhancing experiment design in biomedical research. They utilized mathematical models to see when cells are going to the tumor stage.

That same year, federal policymakers announced the Materials Genome Initiative, which aims to accelerate the discovery of new advanced materials by combining the use of computational and experimental tools along with digital data. Over the last eight years, nationwide, much time, money and resources have been invested in this effort.
Qian and Dougherty turned their focus to materials science problems in 2013. The team started working on optimal design problems two years ago, initially collaborating with Drs. Turab Lookman and Prasanna Balachandran from Los Alamos National Laboratory. Current paradigms are typically centered around the idea of exploring the materials space through experimentation or computation and their approach showed that there are more efficient ways of discovering materials.

"While other people were focusing on the generation and analysis of huge amounts of data, we realized that the best way forward was to focus on experiment design—how to explore the vast domain of possible materials and increase our chances of success by choosing materials with a goal, target property, or response in mind," said Talapatra.

The research is published in the paper, "Autonomous efficient experiment design for materials discovery with Bayesian model averaging", in Physical Review Materials.


Provided by Texas A&M University

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.