Lincoln To
CS 35L – Lab 05
Assignment 10

A Review of AI and Machine Learning in the Discovery of New Materials

I. Introduction

Before discussing the applications of AI and Machine Learning in the discovery of new materials, a brief introduction to the field of Materials Science is useful to provide context to this review. The field of Materials Science and Engineering focuses on the study of the interrelationships between the structure and properties of materials. The following figure illustrates this, indicating that the processing of a material will then affect its underlying structure (on the micro and macro scale), which will define the material's properties, and thus describe the performance of the material.

Figure 1- The four components of the discipline of materials science and engineering and their interrelationship

The field of Materials Science and Engineering is inherently interdisciplinary, and has wide reaching applications throughout all fields and industries. Advances in this field have made further research into fields like efficient solar cells and quantum computing possible, thus, the ability to apply AI and machine learning to accelerate the process of materials research and design is opens up a diverse array of new possibilities throughout a multitude of industries. Kebotix, a startup that combines both robotics and AI to run a self-driving lab, takes advantage of this ability. Their lab, as profiled in ACM TechNews, uses this self-driving lab to accelerate the process of discovering new advanced materials for use in industry.

II. AI and Materials Design

The current approach to materials research has 7 discrete stages, as seen in Figure 2. Although the details of these stages are unimportant for this review, it is important to note that the current approach essentially amounts to a blind search for new materials. It is limited by time, resources, experimental conditions, and theoretical foundations; the time frame from initial research to first use of a material is approximately 10-20 years. Obviously, this is a significant bottleneck in the advancement of various fields of science, technology, and engineering, as many of these fields rely on new materials in order to stay on the bleeding edge of innovation.

Figure 2 - The traditional method of finding new materials

Alternatively, the applications of AI and machine learning to discover and design new materials significantly accelerates this process, lowering the time frame to as much as 18 months. In following the classical definition of machine learning, this approach uses the Performance, Task, Experience (PTE) model to apply to the paradigm of Goal + Sample + Algorithm = Model in order to attempt to discover novel and useful materials. In this case, the goal is the given

problem to be solved. The sample is the population selected for study, and involves data preprocessing such as data cleaning and feature engineering. The algorithms applied can be divided into the machine learning algorithms and model optimization algorithms. The machine learning algorithms commonly used for the research of new materials are support vector machines (SVM), artificial neural networks (ANN), and decision trees (DT). Similarly, the algorithms commonly used for model optimization are genetic algorithms (GA), simulated annealing algorithms (SAA), and particle swarm optimization (PSO).

The methodology of machine learning in materials discovery and design involves 3 main steps: sample construction, model building, and model evaluation. Step 1, sample construction, involves obtaining the sample data from computational simulations and experimental measurements and is followed by data preprocessing. Step 2, model building, is used to model the complex relationships between conditional factors and target attributes. This is where machine learning methods are the most valuable. Finally, step 3, model evaluation, is used to evaluate the model built to confirm that it is effective on new and unknown datasets.

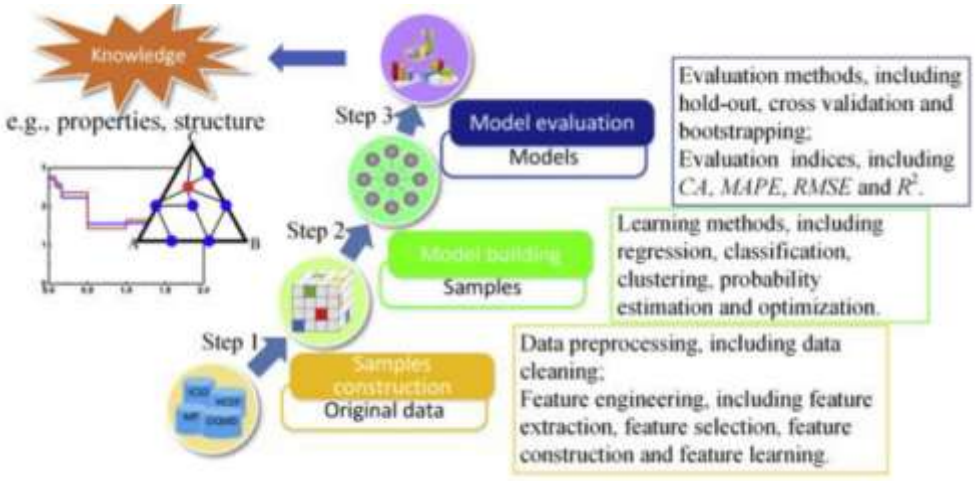

Figure 3 - The methodology of machine learning in materials research and design

With all this noted, machine learning in the field of materials science has a wide variety of practical applications. Currently, machine learning is used primarily for one of two things: property prediction and new materials discovery. In property prediction, regression analysis is commonly used to attempt to predict the macro and microscopic properties of a material. On the other hand, new materials discovery utilizes probabilistic models to screen combinations of structures and components and attempts to select a material with good performance from a candidate set. Other applications include process optimization, in which process parameters must be effectively designed for the synthesis of materials. With this wide variety of applications, it is important to note that selecting the appropriate machine learning algorithms varies depending on the specific application.

All of this leads into the discussion of Kebotix's self-driving lab, the subject of which inspired this review. As mentioned previously, their lab combines both AI and robotics to accelerate the research of new materials. They utilize robotics in order to automate sample

preparation and materials property measurements in order to obtain the raw data for later processing. From there, AI is used to analyze these experimental results and formulate new hypotheses on possible new materials to synthesize and test. More specifically, their lab feeds molecular models of compounds with desirable properties into a neural network, which then learns a statistical representation of these models. Their AI then comes up with new examples that fit this model, which are then synthesized and tested. This process repeats many times, allowing for the rapid iteration of synthesizing and testing new compounds.

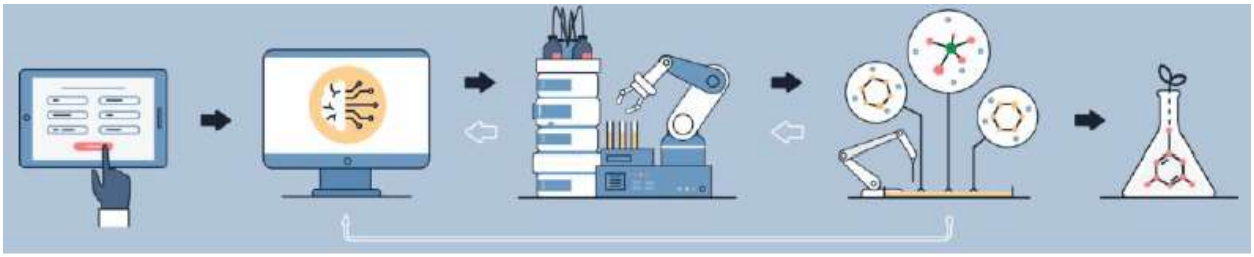

Figure 4 - Schematic of Kebotix's approach to materials research

III. Conclusion

In conclusion, we can see that machine learning has extensive applications in the field of Materials Science and Engineering, particularly in terms of new materials discovery and material property prediction. The general approach is to take classical machine learning methodologies, and apply algorithms such as artificial neural networks, support vector machines, or decision trees, along with knowledge of the field of materials science in order to generate a model that can effectively facilitate these processes of discovery or prediction. This methodology is already been introduced in real world applications as seen in Kebotix's self-driving lab, with the potential to make a dramatic shift in the timeline of discovering useful and novel compounds to advance many other fields of science, technology, and engineering.

Works Cited

1. Knight, Will. "A Robot Scientist Will Dream up New Materials to Advance Computing and Fight Pollution." *MIT Technology Review*, MIT Technology Review, 9 Nov. 2018, www.technologyreview.com/s/612388/a-robot-scientist-will-dream-up-new-materials-to-advance-computing-and-fight-pollution/.
2. Callister, William D., and David G. Rethwisch. *Materials Science and Engineering: An Introduction*. Wiley, 2018.
3. Liu, Yue, et al. "Materials Discovery and Design Using Machine Learning." *Journal of Materiomics*, vol. 3, no. 3, 2017, pp. 159–177., doi:10.1016/j.jmat.2017.08.002.