THE JOURNAL OF PHYSICAL CHEMISTRY LETTERS
A JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

pubs.acs.org/JPCL

Letter

Accelerated Discovery of Zeolite Structures with Superior Mechanical Properties via Active Learning

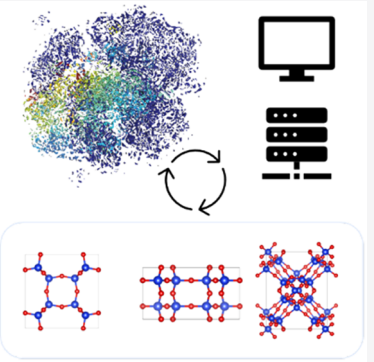
Namjung Kim* and Kyoungmin Min*

Cite This: *J. Phys. Chem. Lett.* 2021, 12, 2334−2339

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: A Bayesian active learning platform is developed for the accelerated discovery of mechanically superior zeolite structures from more than half a million hypothetical candidates. An initial database containing the mechanical properties of synthesizable zeolites is trained to develop the machine learning regression model. Then, a Bayesian optimization scheme is implemented to identify zeolites with potentially excellent mechanical properties. The newly accumulated database consists of 871 labeled structures, and the uncertainty of the predictive model is reduced by 40% and 58% for the bulk and shear moduli, respectively. The model convergence shows that no further improvement occurs after the 10th iteration of optimizations. The proposed platform is able to discover 23 new zeolite structures that have unprecedented shear moduli; in one case, the shear modulus (127.81 GPa) is 250% higher than the previous data set. The proposed platform accelerates the material discovery process while maximizing computational efficiency and enhancing the predictive accuracy.



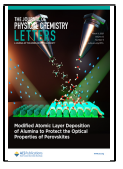
Siliceous zeolites, crystalline metastable phases of SiO_2 with high surface-to-volume and mass ratios, exhibit promising chemical and physical properties that are adjustable depending on the requirements of their applications, such as the adsorption of polar materials, separation based on their porous nature, and ion exchange.[1,2] In addition, they are widely employed in industry as efficient catalysts for biomass transformation, electrode and membrane materials of fuel cells, and adsorption materials for gases such as CO_2.[3−6]

Zeolite frameworks have various intrinsic structural characteristics, including pore configuration, dimensionality, and atomic bond connectivity, leading to immense structural variety. Therefore, the number of possible hypothetical structures of zeolite frameworks is more than half a million, according to the Predicted Crystallography Open Database (PCOD).[7,8] However, the number of experimentally validated structures is approximately 200, as determined by the International Zeolite Association (IZA).[9] This is because of the difficulties in synthesizing zeolites with optimized structural, chemical, and physical properties for target applications.[10] Furthermore, predicting the properties of these zeolites before their synthesis is close to impossible because of their vast structural features.

In this respect, several studies have been attempted, which mainly focused on searching for zeolite structures with ideal properties using machine learning and deep learning. For example, to obtain a complete map of the elastic properties of hypothetical zeolites, machine learning based on a regression model was implemented by training a data set previously constructed from the results of density functional theory (DFT) calculations and zeolite-specific features.[11] This work was further extended to discover auxetic zeolite structures via similar approaches with a greater amount of data.[12] Furthermore, to generate zeolite structures for the adsorption of methane, the inverse design approach was implemented with the aid of generative adversarial networks, which can greatly reduce the number of trials-and-errors in the material design process.[13] Natural language processing is another great tool which can be applied to analyze zeolite synthesis processes and parameters by extracting synthesis information and trends from numerous published journal articles.[10] This information was further used to develop a machine learning regression model for predicting the zeolite framework density as an indicator of structural stability. In addition, combined with a computational geometry approach, machine learning was also implemented to accurately predict and classify the topological types of zeolite crystals.[14,15] For predicting and finding chemically synthesizable zeolite structures, a number of research studies have been conducted with aid of data mining tools, machine learning algorithms, and automated data crawling platforms for the published literature.[16−18]

The current study is primarily focused on finding zeolite frameworks with superior mechanical properties by implementing Bayesian active learning, which includes machine learning and Bayesian optimization. Since zeolites employed in gas adsorption or selective ion exchange membranes are expected to

Received: January 30, 2021
Accepted: February 23, 2021
Published: March 2, 2021



https://dx.doi.org/10.1021/acs.jpclett.1c00339
J. Phys. Chem. Lett. 2021, 12, 2334−2339

withstand various loading conditions, high mechanical properties are critical design requirements.[11,12,19−21] As mentioned, a number of relevant studies for predicting the elastic properties of zeolite structures are available, but additional critical steps are still required, including (1) validation of the predicted properties of the structures (around half a million) and (2) optimization of the developed regression model by training it with additional databases to minimize the prediction uncertainty.

In this study, Bayesian active learning was implemented to search efficiently for zeolite structures with superior mechanical properties. First, to construct the initial database, the mechanical properties of the experimentally validated zeolite frameworks in the IZA database were determined using DFT calculations. Then, based on this database, a machine learning regression model was trained and the elastic properties of the remaining zeolite structures in the PCOD database were predicted. The Bayesian optimization method was then followed to choose the materials for further validation to achieve optimal performance in terms of finding structures with better mechanical properties and to improve the prediction accuracy with minimal trials.

To compute the mechanical properties of the zeolite, DFT calculations with the projector augmented wave (PAW) method in the Vienna *ab initio* simulation package (VASP) were implemented.[22,23] The generalized gradient approximation (GGA) by Perdue, Burke, and Ernzerhof (PBE) was used for the exchange-correlation functional.[24] A plane-wave cutoff of 650 eV was used to provide better convergence. An energy convergence criterion of 10^{-6} eV and a k-point density of 1000 per reciprocal atom were used. The elastic tensor matrix was constructed based on three displacements (central difference) with a step size of 0.015 Å. The elastic properties were obtained from this tensor matrix (shear modulus, bulk modulus, Poisson's ratio, and elastic anisotropy), and the relevant equations are shown in the Supporting Information (SI).

To construct the predictive model for the bulk and shear moduli of the zeolites in the database, we used a gradient boosting-based algorithm, called LightGBM.[25] The LightGBM algorithm utilizes gradient-based one-side sampling and exclusive feature bundling to increase its accuracy and efficiency. Therefore, it is widely used for classical machine learning tasks with large-scale data sets.[26,27] This algorithm was chosen over the other models, such as decision tree (DT), support vector machine (SVM), Gaussian process regressor (GPR), and gradient boosting regressor (GBR), because LightGBM is effective and accurate, especially with large-scale data and features, which corresponds to our problem.

To maximize the prediction accuracy, we optimized the hyper-parameters of LightGBM using RandomizedSearchCV in the scikit-learn library.[28] Therefore, the best hyper-parameters were found with relatively low computational resources. Training and testing used 80% and 20% of the total data set, respectively. The 5-fold cross-validation method was also performed to split the training set and the validation set. Stratified sampling was utilized to sample each data set uniformly over the entire domain. The number of bins for the stratified sampling was set to 8, and the range of each bin was determined based on the minimum and maximum values of the data set.

To select the ideal regression model for the current study, the predictive accuracies of six different machine learning (ML)-based regression models, such as DT, boosting tree ensemble (BTE), SVM, GPR, GBR, and LightGBM, were compared. The minimum leaf size of DT was 4 and the polynomial degree of SVM was set to 2. The numbers of the minimum leaf and BTE learner were 8 and 30, respectively. The learning rate of BTE was set to 0.1. Marten 5/2 was used as the covariance function of GPR. The hyper-parameter values of GBR were obtained from the study by Evans and Coudert[11] in which the best prediction accuracy using the IZA data set was reported.

Active learning is a strategy of learning algorithms that iteratively finds the candidates for labeling, thus maximizing the model accuracy with minimum computational cost.[29] Among the various methods that have been proposed for active learning, Bayesian optimization was chosen to reduce the uncertainty and find the best cases of the predictive model because it is well suited for cases where the cost of obtaining new labels is expensive.[30] In this study, expected improvement (EI) was used as the acquisition function for Bayesian optimization.[31] The EI consists of two components that control the exploitation and exploration, and the coefficient for this trade-off was set to 0.01.

The purpose of the current work is to combine a machine learning predictive model with DFT calculations to find the mechanically superior zeolite structures in an existing database of hypothetical zeolites (PCOD) in which the number of hypothetical structures is approximately 590000. Since brute-force calculation of all of the structures is too expensive and highly ineffective, it is critical that an optimization scheme should be introduced which can greatly reduce the number of trials-and-errors. The specific goals of this study are the following: (1) to suggest zeolite structures with shear (G) and bulk (B) moduli that are larger than those reported in the IZA database (G $\sim$ 55 GPa, B $\sim$ 110 GPa); (2) to reduce the prediction uncertainty; and (3) to confirm convergence of the elastic properties of the unexplored zeolite structures in PCOD.

In this respect, to construct the initial database, 121 structures with previously calculated elastic properties[11,20] were chosen from the IZA database, but their elastic properties were recalculated with the current calculation scheme. This is because a previous study implemented the B3LYP hybrid functional with localized basis sets, whereas our work employed the PAW-based DFT method with a different type of pseudopotential. Subsequently, the results between the two different methods were compared, which revealed no distinct difference, as shown in Figure S1, thus validating the reliability of the current approach.

The Bayesian active learning process is presented in Figure 1, and the details of each step are as follows:

(1) Based on the calculated mechanical properties of the zeolite structures in the IZA database, a machine learning model was constructed to predict the shear and bulk moduli. Previously developed and validated zeolite descriptors were adopted.[11]

(2) The constructed regression model was then applied to predict the elastic properties of the remaining zeolite structures in PCOD. To quantify the prediction uncertainty, 50 regression models were generated by changing the composition of the training set.

(3) The results from the uncertainty analysis (predicted mean and standard deviation values) were provided to the Bayesian optimization method to identify the materials. Fifty structures with larger values of EI were chosen, and their elastic properties were calculated using DFT. Zeolite structures, in PCOD, with more than 66 atoms were not considered during the optimization for computational efficiency. During the DFT calculations, structures with negative shear and bulk moduli, structures with invalid moduli, and structures that did not converge during energy calculation were removed from the

https://dx.doi.org/10.1021/acs.jpclett.1c00339
J. Phys. Chem. Lett. 2021, 12, 2334−2339

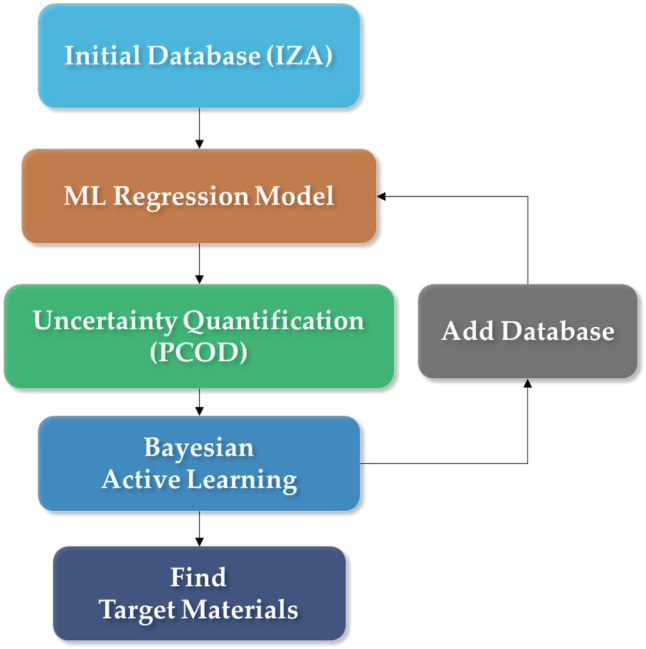


Figure 1. Bayesian active learning platform to search for mechanically superior zeolite structures.

updating list. The updated data set was trained again with the regression model.

(4) Steps (2) and (3) were repeated 20 times to reduce the prediction uncertainty and to discover zeolites with superior mechanical properties.

The bulk modulus determined by DFT calculations was plotted with those determined by the ML-based regression models, as displayed in Figures S2 and S3. The results indicate that, among the six different regression models, LightGBM is the best regression model for predicting the bulk and shear moduli of zeolites in terms of accuracy and repeatability; hence, it was used for the rest of the study.

After confirming the accuracy of the predictive model with the limited amount of data, active learning was performed to enhance the model accuracy through the gradual addition of labeled data by an iterative learning process. The convergence and accuracy of the predictive model were monitored as the iterations proceeded. Figure 2a–c show the representative predictions of the testing data set for the bulk and shear moduli at the 0th, 10th, and 20th iteration. First, for the bulk and shear moduli at the zeroth iteration, the mean absolute errors (MAEs) were 3.72 GPa (bulk) and 2.16 GPa (shear) and the R^2 scores were 0.947 (bulk) and 0.885 (shear). As the iterations proceeded, the MAEs slightly increased to 7.59 GPa (bulk) and 2.98 GPa (shear) at the 10th iteration and plateaued around 6.75 GPa (bulk) and 2.69 GPa (shear) at the 20th iteration. The R^2 score at the 10th and 20th iteration showed a similar trend as

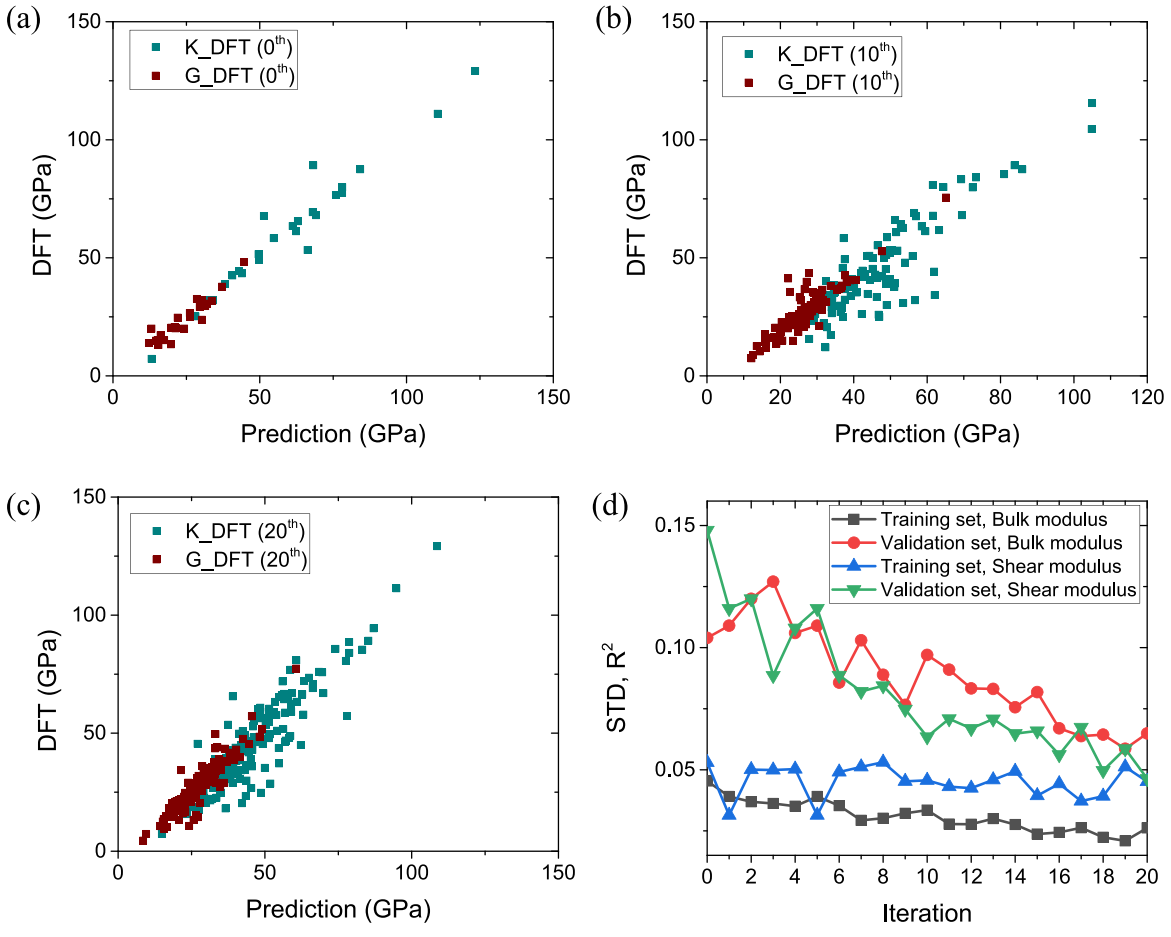


Figure 2. (a–c) Representative predictions of the testing data set of the bulk and shear moduli at the 0th, 10th, and 20th iteration. (d) Standard deviation (R^2 score) of the predictive models as iteration proceeds.

https://dx.doi.org/10.1021/acs.jpclett.1c00339
J. Phys. Chem. Lett. 2021, 12, 2334–2339

that of the MAE values. This indicated that improvement of the predictive model was saturated at high accuracy. Notably, although more data points were added, there was no sign of further improvement in predictive accuracy. Although this is rather unexpected, this can be justified as follows: (1) The actual number of structures (~900) with elastic properties is extremely small compared with the unexplored region (~590000); hence, a certain limitation in terms of accuracy may exist for the predictive model because of the asymmetrical size of the labeled and unlabeled data sets. (2) It is more important to validate the current framework for finding superior materials despite this uncertainty and reduce the prediction uncertainty significantly where further optimization is not necessary.

In this regard, the convergence of the predictive model was further validated by observing the standard deviation (R^2 value) for predictions of the bulk and shear moduli, as shown in Figure 2d. The R^2 score represents the quantitative deviation of the predictive values among the 50 regression models given the training and validation data sets. As the iterations proceeded, the standard deviation of the predictions using the validation set reduced and reached certain values (from approximately 0.15 to 0.05 for both cases). This value was almost comparable to that obtained from the training set, confirming that the predictive uncertainty was largely reduced. In other words, the 50 regression models produced consistent predictions, and improvement of the predictive model was saturated at a certain limit.

After confirming the convergence of the predictive model based on the labeled data, the predictive model with the ~590000 unlabeled data set was also studied to measure the convergence of the predictive model. This is the most critical step in the validation process because convergence should be satisfied to further utilize the current predictive model (too many unlabeled data remained to be validated with DFT calculations). Figure 3 displays the standard deviation (R^2 score)

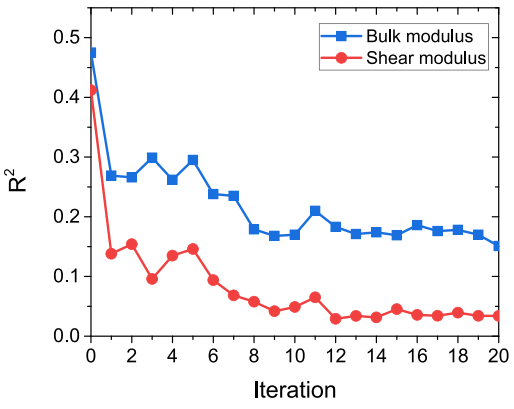


Figure 3. Standard deviation (R^2 score) of the predictive model with the unlabeled data set for the bulk and shear moduli.

of the predictive model as the iterations proceeded with the addition of the unlabeled data set. The standard deviation of the predictive model also converged to a certain value (approximately 0.15 for the bulk modulus and 0.04 for the shear modulus) after the 15th iteration, similar to the training and validation data sets, indicating convergence of the predictive model with respect to the unlabeled data set.

On the basis of this result, we confirmed that the convergence of our predictive model and the 20 iterations of active learning were sufficient to induce distinctive and finalized improvement of our predictive model. In addition, this result indicates that the addition of more data to certain iterations increases the predictive accuracy and reduces the uncertainty in the prediction. Furthermore, small labeled data sets and the absence of data addition limit the accuracy improvement of the predictive model. At the final iteration, our database has 871 labeled zeolite structures, which is seven times larger than the initial IZA data set. Because of the large amount of data and the efficient sampling strategy, we believe that our predictive model produces a refined estimation for predicting the bulk and shear moduli of zeolites compared with previous studies.

In order to elucidate the descriptors used in this study and their correlation changes before and after adding more data sets, the interdependencies of the implemented descriptors were obtained, as shown in Figure S5, from the IZA data set and from the data set after the 20th iteration. The results indicate that the dependencies of the most of pairwise descriptors are affected as the labeled data are added. Hence, including all the descriptors for the regression model is recommended for the current model in order to obtain all the unexpected effects from the entire descriptors. Furthermore, the interpretation of ML models is obtained by analyzing the partial dependence plots, as shown in Figure S6 and Table S2. It indicates that the information from three categories (local, structural, and porosity) of zeolite descriptors are equally important and necessary to predict the bulk and shear moduli of a zeolite structure.

Upon achieving model convergence, the zeolite structures found during active learning were examined for superior mechanical properties. Figure 4a and 4b respectively show the calculated bulk and shear moduli of the queried points as iteration proceeds. Each point represents the bulk and shear moduli of the zeolite structure calculated by DFT, which was queried by the Bayesian optimization algorithm. The bulk and shear moduli at Iteration 0 were obtained from the zeolite structures in the IZA database. Figure S4 shows the continuous densities of the bulk and shear moduli estimated by kernel density estimation (KDE). For the bulk modulus, all the queried points existed within the range of the IZA data set (7.22−129.14 GPa), indicating that the IZA database covered most of the feasible range of the bulk modulus. However, for the shear modulus, our Bayesian active learning platform discovered 23 new zeolite structures with shear moduli higher than those in the IZA database (5.61−48.16 GPa), as listed in Table S1. The maximum discovered shear modulus was 127.81 GPa, which is more than 2.5 times higher than the maximum value in the IZA data set. The shear modulus is a measure of the resistance against shear loading; a zeolite with a 2.5 times higher shear modulus is able to resist a 250% higher shear loading, which can satisfy the requirement of outstanding mechanical performance while maintaining chemical advantages.[32]

In this work, a Bayesian active learning platform was developed for the accelerated discovery of mechanically superior zeolite structures. An initial database containing the mechanical properties of experimentally synthesizable zeolite structures was constructed to train the machine learning regression model. Then, Bayesian optimization was applied to the hypothetical zeolite structures (more than half a million in the PCOD) to identify materials with potentially excellent mechanical properties, which significantly reduces the computational cost. A rigorous comparison of the results indicates that the selected ML-based regression model, LightGBM, is the best model for the active learning platform in terms of accuracy and repeatability. The new database contains a total of 871 labeled

https://dx.doi.org/10.1021/acs.jpclett.1c00339
J. Phys. Chem. Lett. 2021, 12, 2334−2339

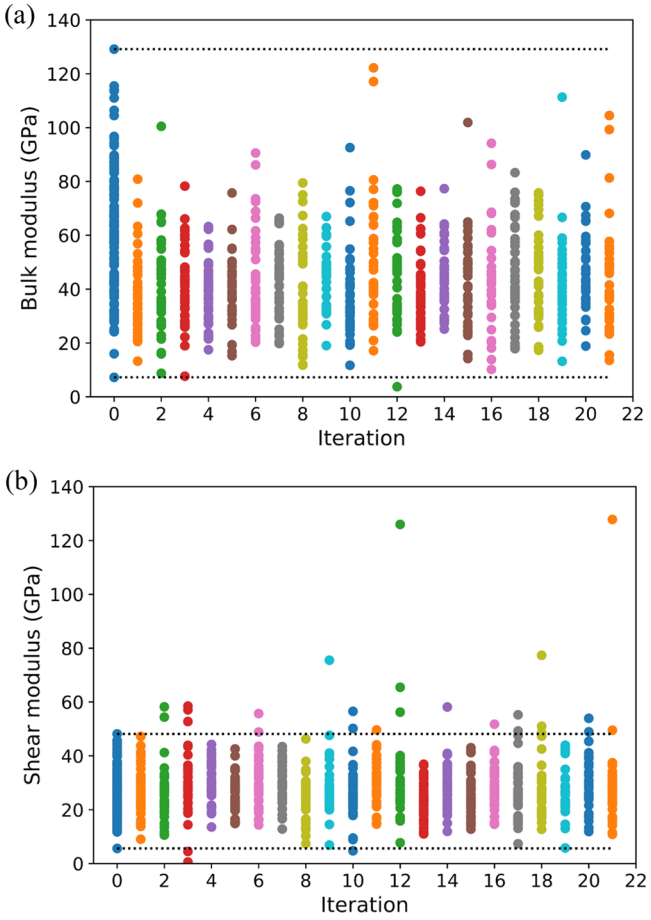


Figure 4. DFT-calculated (a) bulk and (b) shear moduli of the queried zeolite structures as Bayesian active learning proceeds. The dotted line represents the minimum and the maximum modulus that the IZA database covers.

zeolite structures, which is seven times larger than the IZA data set.

Because of the strategic addition of labeled data and the well-trained regression model, the proposed platform was able to reduce the uncertainty of the prediction for the unlabeled data set by 40% and 58% for the bulk and shear moduli, respectively. Importantly, this was achieved by labeling only 0.15% of the total data. The proposed platform offers a universal tool to accelerate the discovery of new material. Moreover, it can be applied to find other materials with unprecedented mechanical properties, such as unusually high Poisson's and thermal expansion ratios. The 23 newly discovered zeolite structures have unprecedented shear moduli, demonstrating the huge potential of using zeolites as functional materials in mechanics as well as in chemistry-related applications.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jpclett.1c00339.

Detailed description of elastic properties derivation, validation of the current method, predictive model accuracy, repeatability of the prediction model, distribution of predicted bulk and shear moduli, and list of superior zeolite structures (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Namjung Kim − *Department of Mechanical Engineering, Gachon University, Seongnam, Gyeonggi-do 13120, Republic of Korea;* orcid.org/0000-0002-2600-5921; Email: namjungk@gachon.ac.kr

Kyoungmin Min − *School of Mechanical Engineering, Soongsil University, Seoul 06978, Republic of Korea;* orcid.org/0000-0002-1041-6005; Email: kmin.min@ssu.ac.kr

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jpclett.1c00339

Notes

Data Availability: Machine learning scripts, the training set, and the test set used in this study can be found at https://github.com/kminmin/zeolitebayesian.

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by a National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIT) (2020R1F1A1073300, 2020R1F1A1066519). This work was supported by the National Supercomputing Center with supercomputing resources and including technical support (KSC-2020-CRE-0012).

■ REFERENCES

(1) Davis, M. E. Ordered Porous Materials for Emerging Applications. *Nature* **2002**, *417* (6891), 813−821.

(2) Martínez, C.; Corma, A. Inorganic Molecular Sieves: Preparation, Modification and Industrial Application in Catalytic Processes. *Coord. Chem. Rev.* **2011**, *255* (13), 1558−1580.

(3) Moliner, M. State of the Art of Lewis Acid-Containing Zeolites: Lessons from Fine Chemistry to New Biomass Transformation Processes. *Dalt. Trans.* **2014**, *43* (11), 4197−4208.

(4) Chen, Z.; Holmberg, B.; Li, W.; Wang, X.; Deng, W.; Munoz, R.; Yan, Y. Nafion/Zeolite Nanocomposite Membrane by in Situ Crystallization for a Direct Methanol Fuel Cell. *Chem. Mater.* **2006**, *18* (24), 5669−5675.

(5) Wu, H.; Zheng, B.; Zheng, X.; Wang, J.; Yuan, W.; Jiang, Z. Surface-Modified Y Zeolite-Filled Chitosan Membrane for Direct Methanol Fuel Cell. *J. Power Sources* **2007**, *173* (2), 842−852.

(6) Siriwardane, R. V.; Shen, M.-S.; Fisher, E. P.; Losch, J. Adsorption of CO2 on Zeolites at Moderate Temperatures. *Energy Fuels* **2005**, *19* (3), 1153−1159.

(7) Earl, D. J.; Deem, M. W. Toward a Database of Hypothetical Zeolite Structures. *Ind. Eng. Chem. Res.* **2006**, *45* (16), 5449−5454.

(8) Pophale, R.; Cheeseman, P. A.; Deem, M. W. A Database of New Zeolite-like Materials. *Phys. Chem. Chem. Phys.* **2011**, *13* (27), 12407−12412.

(9) IZA Structure Commission. Database of Zeolite Structures. http://www.iza-structure.org/databases/.

(10) Jensen, Z.; Kim, E.; Kwon, S.; Gani, T. Z. H.; Román-Leshkov, Y.; Moliner, M.; Corma, A.; Olivetti, E. A Machine Learning Approach to Zeolite Synthesis Enabled by Automatic Literature Data Extraction. *ACS Cent. Sci.* **2019**, *5* (5), 892−899.

(11) Evans, J. D.; Coudert, F.-X. Predicting the Mechanical Properties of Zeolite Frameworks by Machine Learning. *Chem. Mater.* **2017**, *29* (18), 7833−7839.

(12) Gaillac, R.; Chibani, S.; Coudert, F.-X. Speeding Up Discovery of Auxetic Zeolite Frameworks by Machine Learning. *Chem. Mater.* **2020**, *32* (6), 2653−2663.

(13) Kim, B.; Lee, S.; Kim, J. Inverse Design of Porous Materials Using Artificial Neural Networks. *Sci. Adv.* **2020**, *6* (1), eaax9324.

(14) Carr, D. A.; Lach-hab, M.; Yang, S.; Vaisman, I. I.; Blaisten-Barojas, E. Machine Learning Approach for Structure-Based Zeolite Classification. *Microporous Mesoporous Mater.* **2009**, *117* (1), 339−349.
(15) Yang, S.; Lach-hab, M.; Vaisman, I. I.; Blaisten-Barojas, E. Identifying Zeolite Frameworks with a Machine Learning Approach. *J. Phys. Chem. C* **2009**, *113* (52), 21721−21725.
(16) Daeyaert, F.; Ye, F.; Deem, M. W. Machine-Learning Approach to the Design of OSDAs for Zeolite Beta. *Proc. Natl. Acad. Sci. U. S. A.* **2019**, *116* (9), 3413−3418.
(17) Moliner, M.; Román-Leshkov, Y.; Corma, A. Machine Learning Applied to Zeolite Synthesis: The Missing Link for Realizing High-Throughput Discovery. *Acc. Chem. Res.* **2019**, *52* (10), 2971−2980.
(18) Muraoka, K.; Sada, Y.; Miyazaki, D.; Chaikittisilp, W.; Okubo, T. Linking Synthesis and Structure Descriptors from a Large Collection of Synthetic Records of Zeolite Materials. *Nat. Commun.* **2019**, *10* (1), 4459.
(19) Ambekar, R. S.; Oliveira, E. F.; Kushwaha, B.; Pal, V.; Machado, L. D.; Sajadi, S. M.; Baughman, R. H.; Ajayan, P. M.; Roy, A. K.; Galvao, D. S.; et al. On the Mechanical Properties of Atomic and 3D Printed Zeolite-Templated Carbon Nanotube. *Networks. Addit. Manuf.* **2021**, *37*, 101628.
(20) Coudert, F.-X. Systematic Investigation of the Mechanical Properties of Pure Silica Zeolites: Stiffness, Anisotropy, and Negative Linear Compressibility. *Phys. Chem. Chem. Phys.* **2013**, *15* (38), 16012−16018.
(21) Tan, J. C.; Bennett, T. D.; Cheetham, A. K. Chemical Structure, Network Topology, and Porosity Effects on the Mechanical Properties of Zeolitic Imidazolate Frameworks. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107* (22), 9938−9943.
(22) Kresse, G.; Furthmüller, J. Efficiency of Ab-Initio Total Energy Calculations for Metals and Semiconductors Using a Plane-Wave Basis Set. *Comput. Mater. Sci.* **1996**, *6* (1), 15−50.
(23) Kresse, G.; Furthmüller, J. Efficient Iterative Schemes for Ab-Initio Total Energy Calculations Using a Plane-Wave Basis Set. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1996**, *54* (16), 11169−11186.
(24) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77* (18), 3865−3868.
(25) Ke, G.; Meng, Q.; Finley, T.; Wang, T.; Chen, W.; Ma, W.; Ye, Q.; Liu, T.-Y. LightGBM: A Highly Efficient Gradient Boosting Decision Tree. In *Advances in Neural Information Processing Systems 30*; Guyon, I., Luxburg, U. V, Bengio, S., Wallach, H., Fergus, R., Vishwanathan, S., Garnett, R., Eds.; Curran Associates, Inc.: 2017; pp 3146−3154.
(26) Wang, D.; Zhang, Y.; Zhao, Y. LightGBM: An Effective MiRNA Classification Method in Breast Cancer Patients. In *Proceedings of the 2017 International Conference on Computational Biology and Bioinformatics*; 2017; pp 7−11.
(27) Sun, X.; Liu, M.; Sima, Z. A Novel Cryptocurrency Price Trend Forecasting Model Based on LightGBM. *Financ. Res. Lett.* **2020**, *32*, 101084.
(28) Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blndel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-Learn: Machine Learning in Python. *J. Mach. Learn. Res. Pref.* **2011**, *39*, 2825−2830.
(29) Müller, W. G. *Collecting Spatial Data: Optimum Design of Experiments for Random Fields*; Springer Science & Business Media: 2007.
(30) Shahriari, B.; Swersky, K.; Wang, Z.; Adams, R. P.; De Freitas, N. Taking the Human out of the Loop: A Review of Bayesian Optimization. *Proc. IEEE* **2016**, *104* (1), 148−175.
(31) Snoek, J.; Larochelle, H.; Adams, R. P. Practical Bayesian Optimization of Machine Learning Algorithms Jasper. *Adv. Neural Inf. Process. Syst.* **2012**, 2951−2959.
(32) Wang, S.; Bai, P.; Sun, M.; Liu, W.; Li, D.; Wu, W.; Yan, W.; Shang, J.; Yu, J. Fabricating Mechanically Robust Binder-Free Structured Zeolites by 3D Printing Coupled with Zeolite Soldering: A Superior Configuration for CO2 Capture. *Adv. Sci.* **2019**, *6* (17), 1901317.

https://dx.doi.org/10.1021/acs.jpclett.1c00339
J. Phys. Chem. Lett. 2021, 12, 2334−2339