ACS APPLIED MATERIALS & INTERFACES

www.acsami.org

Research Article

Accurate Prediction of Voltage of Battery Electrode Materials Using Attention-Based Graph Neural Networks

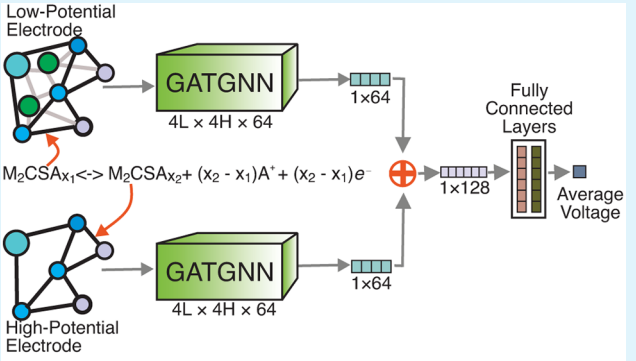
Steph-Yves Louis, Edirisuriya M. Dilanga Siriwardane, Rajendra P. Joshi,* Sadman Sadeed Omee, Neeraj Kumar, and Jianjun Hu*

Cite This: *ACS Appl. Mater. Interfaces* 2022, 14, 26587−26594

Read Online

ACCESS | Metrics & More | Article Recommendations

ABSTRACT: Performing first-principles calculations to discover electrodes' properties in the large chemical space is a challenging task. While machine learning (ML) has been applied to effectively accelerate those discoveries, most of the applied methods ignore the materials' spatial information and only use predefined features: based only on chemical compositions. We propose two attention-based graph convolutional neural network techniques to learn the average voltage of electrodes. Our proposed methods, which combine both atomic composition and atomic coordinates in 3D-space, improve the accuracy in voltage prediction significantly when compared to composition-based ML models. The first model directly learns the chemical reaction of electrodes and metal ions to predict their average voltage, whereas the second model combines electrodes' ML predicted formation energy (E_{form}) to compute their average voltage. Our E_{form}-based model demonstrates improved accuracy in transferability from our subset of learned Li ions to Na ions. Moreover, we predicted the theoretical voltage of 10 Na_xMPO_4F (M = Ti, Cr, Fe, Cu, Mn, Co, and Ni) fluorophosphate battery frameworks, which are unavailable in the Material Project database. It could be shown that we can expect average voltages higher than 3.1 V from those Na battery frameworks except from the $NaTiPO_4F$ and $TiPO_4F$ pair of electrodes, which offer an average voltage of 1.32 V.



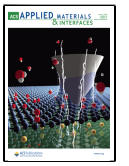
KEYWORDS: *electrodes, graph neural network, voltage, deep learning, formation energy*

1. INTRODUCTION

Batteries are the dominant source of energy for diverse applications and the main workhorse for portable electronics.[1,2] Common examples where batteries are increasingly adopted are electric vehicles and grid energy storage.[3,4] Besides their wide use, there are still big interests in improving these batteries' performance for more improved reliability in devices demanding large energy density. But to develop next-generation batteries, accurate and efficient exploration of large chemical space is necessary, and predicting their performance represents the first step toward this goal. Traditionally, the exploration of these batteries' or electrodes' properties was done by using time-consuming physics-based simulations[5] and/or by using resource-intensive experiments.[6] Mainly, examining each material in the large chemical space while searching for robust electrodes imposes great difficulties with these traditional methods. Hence, machine learning (ML) has been used as an alternative for their impressive performance and are increasingly adopted in the battery community for predicting the performance metrics of battery components including intercalation potentials or voltage.[2,7−14]

Recently, multiple ML-based approaches to predict the voltage of electrodes materials have been used.[8,15,16] Such approaches include models based on ML potentials and also other simple deep neural network, which are the most accurate. ML-potential-based models are trained to learn the potential energy surface of solids by using data from physics-based simulations.[8] Though promising, the accuracy of these ML potentials is limited to particular types of materials with specific atomic compositions and have weak transferability. Another important limitation to these ML potentials is that generating the data set for their model from density functional theory (DFT), for each composition space, is extremely challenging. For instance, Viswanathan and Houchins adapted ML potentials to predict voltage, only for C-based electrodes, using a relatively small DFT based data set.[8] While the

Received: January 2, 2022
Accepted: May 19, 2022
Published: June 6, 2022



accuracy of Viswanathan's model compares to that of DFT simulations for Li-graphite-based electrodes, its transferability to Na-graphite or K-graphite is impractical. This impracticability results from the fact that the Li-graphite-specific ML potentials model performs poorly on other types of electrodes. Because such data do not exist for each material for other metal-ion batteries, ML models that will work well with diverse metal-ion batteries and transfer equally well are necessary.

For other models including the deep neural network, ML models are trained on nearly 5000 electrodes materials from the materials project database with density functional theory calculated voltage as target.[15,16] Nonetheless, these models have poorer predictive ability than other ML models used for properties of solids.[17] This poor performance can be attributed to both the small size of the data set and chemical diversity within such a small data set for the different metal-ion batteries. Also, more advanced ML methods, such as graph-based ML approaches, had not been applied yet in the domain of electrode materials. The majority of existing literature mostly employs predefined calculated features as input information for the materials. Solely relying on these precalculated atomic composition features implies that the model ignores any 3D structural environment information within the materials and thus misses the key attributes used in reference physics-based simulations.[15,16] However, we anticipate that combining both chemical composition and some 3D structural information inside of a more complex deep learning method may lead to a significantly more accurate voltage prediction, as it has been shown for other crystals' properties predictions.[17−19]

In this work, we adapt a graph-convolutional neural network that learns the robust representation of electrode materials from the atom types and corresponding 3D coordinates only. We show that our method provides gradual improvement upon existing work. Our proposed method shows great transferability to new metal-ion battery chemistry as it outperforms all related published works on voltage prediction. Our voltage prediction solution includes two techniques. The first consists of a model that learns the chemical reaction of input electrodes and outputs their average voltage. The other technique involves using a trained model that predicts E_{form} of individual electrodes. These electrodes' E_{form} are subsequently used within our derived formula to output the voltage prediction. Comparing our performance to those found in the literature, we show that calculating the voltage from formation energy of electrodes is an ideal way to predict voltage in the scenario where there are limited data to train ML models for intercalation reactions.

2. METHODS

2.1. Data. DFT computed voltages and structures of electrodes materials for 4402 battery systems were collected from the Material Project (MP) database using Pymatgen Materials Genomes (pymatgen).[20] The distribution of the data set, which consists of 10 different metal ion (Cs, Y, Zn, Ca, Li, Mg, Na, K, Al, and Rb) batteries, is shown in Figure 1.[21] Because of its high popularity as a charge carrier, Li (2291) has the highest number of battery systems. The other battery frameworks include Ca-based systems (484), Mg (393), Na (328), Zn (385), Rb (50), and Cs (39)-based electrodes. These electrode distributions are displayed in Figure 1.

To develop the two proposed methods below described, we use the previously mentioned data set of 4402 electrodes for the chemical reaction-based model and another data set of about 60000 materials

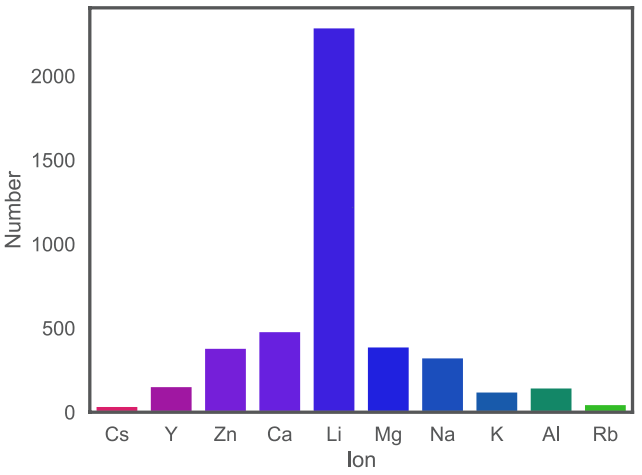


Figure 1. Distribution of the numbers of battery frameworks collected from the MP database for each metal ion.

downloaded from the MP database for E_{form}-based model.[22] The two data sets are mutually exclusive of their materials.

2.2. Voltage. Understanding the chemical reaction of an intercalation battery framework is essential to learn the voltage of the corresponding system. As an example, the intercalation reaction of the hypothetical layered material $A_{x_1}B_\alpha C_\beta D\gamma$ and A ions to form $A_{x_2}B_\alpha C_\beta D\gamma$ can be represented as follows:

$$A_{x_1}B_\alpha C_\beta D\gamma + (x_2 - x_1)A^{z+} + z(x_2 - x_1)e^- \leftrightarrow A_{x_2}B_\alpha C_\beta D\gamma \quad (1)$$

The electrode on the left-hand side, which reacts with the cation, exhibits a higher potential than the electrode on the right-hand side. Thus, we label the left-hand and right-hand side electrodes as high-potential and low-potential electrodes, respectively. To calculate the voltage in DFT, we estimate the Gibbs free energy of individual electrodes defined as $G = \Delta E + P\Delta V' - T\Delta S$, where ΔE is the internal energy change, P is the pressure, $\Delta V'$ is the volume change, T is the temperature, and ΔS is the entropy difference of the system. However, $P\Delta V' \approx 10^{-5}$ eV and $T\Delta S \approx 25$ meV at room temperature. Therefore, by neglecting those two terms, we can calculate the voltage (V) by only considering the internal energy change as shown in eq 2. Here, the terms $E[\eta_{x_i}]$ $(i = 1, 2)$ are the total energy of the chemical formula (η) with x_1 and x_2 contents of the ion (A), and z is the valency of the intercalating metal ion. For instance, $\eta_{x_1} = A_{x_1}B_\alpha C_\beta D\gamma$ and $\eta_{x_2} = A_{x_2}B_\alpha C_\beta D\gamma$ for the chemical reaction of eq 1. For the ions mentioned in Figure 1, $z = 1$ for Li, Na, K, Rb, and Cs, $z = 2$ for Ca, Mg, and Zn, and $z = 3$ for Al and Y.

$$V \approx \frac{1}{z(x_2 - x_1)e}(E[\eta_{x_1}] - E[\eta_{x_2}] + (x_2 - x_1)E(A)) \quad (2)$$

$$E_{form_{x_i}} = E[\eta_{x_i}] - x_1E(A) - \alpha E(B) - \beta E(C) - \gamma E(D) \quad (3)$$

The E_{form} per unit formula of $A_{x_1}B_\alpha C_\beta D\gamma$ electrode is given by eq 3, where E(A), E(B), E(C), and E(D) represent the energy of each atom in their bulk phase. By computing the difference between formation energies of high and low electrodes, we can show that voltage can be determined using eq 4.

$$V \approx \frac{1}{z(x_2 - x_1)e}(E_{form_{x_1}} - E_{form_{x_2}}) \quad (4)$$

However, we can calculate the voltage in two different ways with ML: (I) by training the ML models directly to learn the voltage and (II) by training ML models to learn the formation energy of involved high- and low-potential electrodes. We compare the performance of the model trained on each case. In the first method, the structures corresponding to high- and low-potential electrodes for an

https://doi.org/10.1021/acsami.2c00029
ACS Appl. Mater. Interfaces 2022, 14, 26587−26594

intercalation reaction are used simultaneously (hence labeled reaction-based model). In contrast, for the second method, the structure of each compound is input separately to first learn the formation energy (labeled E_{form}-based model). Once the formation energy is predicted, the corresponding voltage can then be calculated by using eq 4.

2.3. Graph Neural Network Architecture. Graph neural networks (GNN) are deep neural networks that have been applied to effectively learn latent features from network or graph data.[23] Our models are based on a subset of these GNNs that adapt the technique of the attention mechanism to GNN.[24−26] Particularly, we adapt the technique of GATGNN introduced by Louis et al.[27] for this study.

The input to our GATGNN adapted models is materials encoded as a graph with nodes representing atoms and edges representing those nodes connections. We encode a material as a graph where each node (atom) connects with the 16 nearest nodes (atoms). Each atom type is then attributed a 92-dimensional vector, and each edge's distance is also encoded as a 41-dimensional vector.[27]

Compared to other GNN applied in the domain of materials, GATGNN learns each atom's contribution both locally (within a local atomic space) and globally (with respect to all atoms in the material).[17,18,28,29] The overall architecture of the GATGNN model is shown in Figure 2. GATGNN first efficiently captures the atoms' local importance through its augmented graph attention layers (AGAT) and then a global attention.

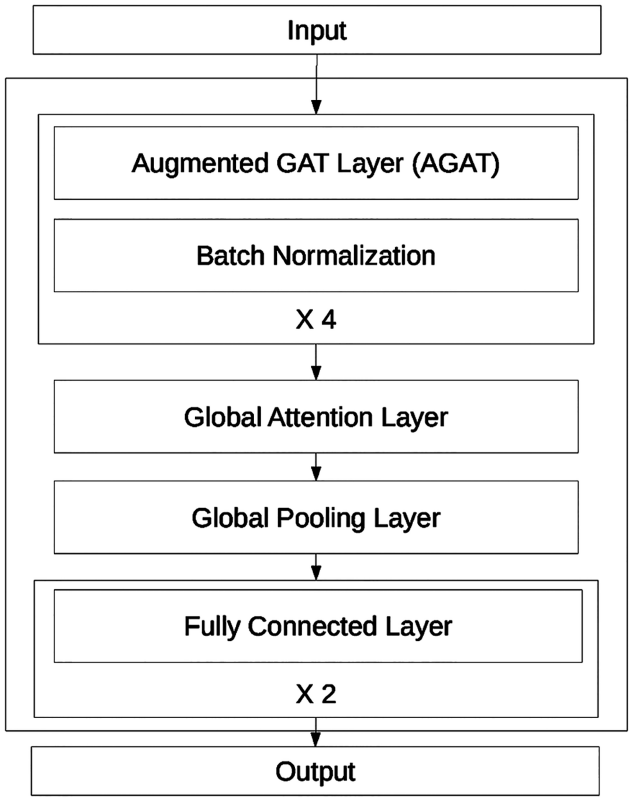


Figure 2. An overview of the GATGNN architecture.

In our research, we adapt the proposed GATGNN to construct our models for learning voltage from materials. For the local attention, our model consist of four AGAT layers of four attention heads, each consisting of 64 neurons. The local soft-attention $\alpha_{i,j}$ between a node i and a neighbor j can be represented as

$$\alpha_{i,j} = \frac{\exp(a_{i,j})}{\sum_{k\in N_i} \exp(a_{i,k})} \tag{5}$$

where N_i denotes the neighborhood of node i and $a_{i,j}$ is the parametrized weight coefficient between nodes i and j, which denotes the importance of node j to node i in eq 5.

Upon learning the local importance of the atoms, we subsequently use a single fully connected layer to learn this global attention value. In our study, the two inputs needed for the global attention layer are an atom's embedding and a material's compositional vector.[27] The global attention g_i can be described as follows:

$$g_i = \frac{(\mathbf{x}_i \| \mathbf{E})\cdot \mathbf{W}}{\sum_{x_c\in \mathbf{x}} (\mathbf{x}_c \| \mathbf{E})\cdot \mathbf{W}} \tag{6}$$

$\mathbf{x} \in \mathbb{R}^F$ denotes a learned embedding, $\mathbf{E}$ denotes a compositional vector of the crystal, $\mathbf{W} \in \mathbb{R}^{1\times(F+|E|)}$ denotes a parametrized matrix, and x_c denotes the learned embedding of any atom c within the crystal in eq 6.

We implement all the components in our proposed approach using deep learning libraries of Pytorch and the library of Pytorch-Geometric.[30,31] The same SmoothL1 loss function is used to train both models.[32]

2.4. Chemical-Reaction-Based Voltage Predictor. Because of their recent advances, ML models have been increasingly adopted for learning the properties of chemical reactions of molecules.[33,34] These models, however, have not been used for reactions involving crystals. Hence, we developed a GNN model that considers the chemical reaction of electrodes and metal ions as the input. Our proposed model is based on the previously existing GATGNN method. Mainly, our reaction-based model consists of two modified GATGNN (GATGNN^R) modules arranged in parallel, which are both followed by series of hidden fully connected layers. Figure 3 illustrates the framework of our proposed model.

To start, a low and a high electrode are input into the low/high dedicated GATGNN^R module which learn from their corresponding electrodes. Following the graph convolutions, the output of both blocks or modules is then concatenated into a 128-dimensional vector to be fed to two fully connected layers. The final predicted average voltage is calculated by learning the chemical interaction of the two electrodes.

Hyperparameters were optimized for all the models used in this work. We train the model for 500 epochs with early stopping using a learning rate of 1×10^{-3}, a weight decay of 5×10^{-3}, and a batch size of 128.

2.5. Formation Energy-Based Voltage Predictor. Our proposed E_{form}-based model is based on an optimized pretrained model of a GATGNN of four layers 128 neurons with four attention heads to model crystals to their E_{form}. As illustrated in Figure 4, the proposed method consists of one GATGNN which independently outputs the E_{form} for the low- and high-potential electrodes. This optimized E_{form}-based GATGNN model was trained for 300 epochs early stopping by using a learning rate of 1×10^{-3}, a weight decay of 5×10^{-3}, and a batch size of 64. For the data set, we split the data into three sets: with 85% used for training, 7.5% for testing, and 7.5% for validation. Following the prediction of the electrodes' E_{form}, their average voltage is subsequently obtained by using eq 4.

2.6. Experiments. To assess the validity of our proposed method, we conducted two different experiments. In the first experiment, we do a 10-fold cross-validation as done in the works by Moses et al.[16] In the second experiment, we apply the holdout test with a data split of 85% used for training, 7.5% for testing, and 7.5% for validation. Notably, we use the first experiment to compare the performance of our proposed method to other voltage works that only use composition. In contrast, we use the second experiment to compare the performance of using the reaction-based model to the E_{form}-based model.

3. RESULTS AND DISCUSSION

To evaluate the performance of the models, we used the metric of mean absolute error (MAE) defined in the equation

https://doi.org/10.1021/acsami.2c00029
ACS Appl. Mater. Interfaces 2022, 14, 26587−26594

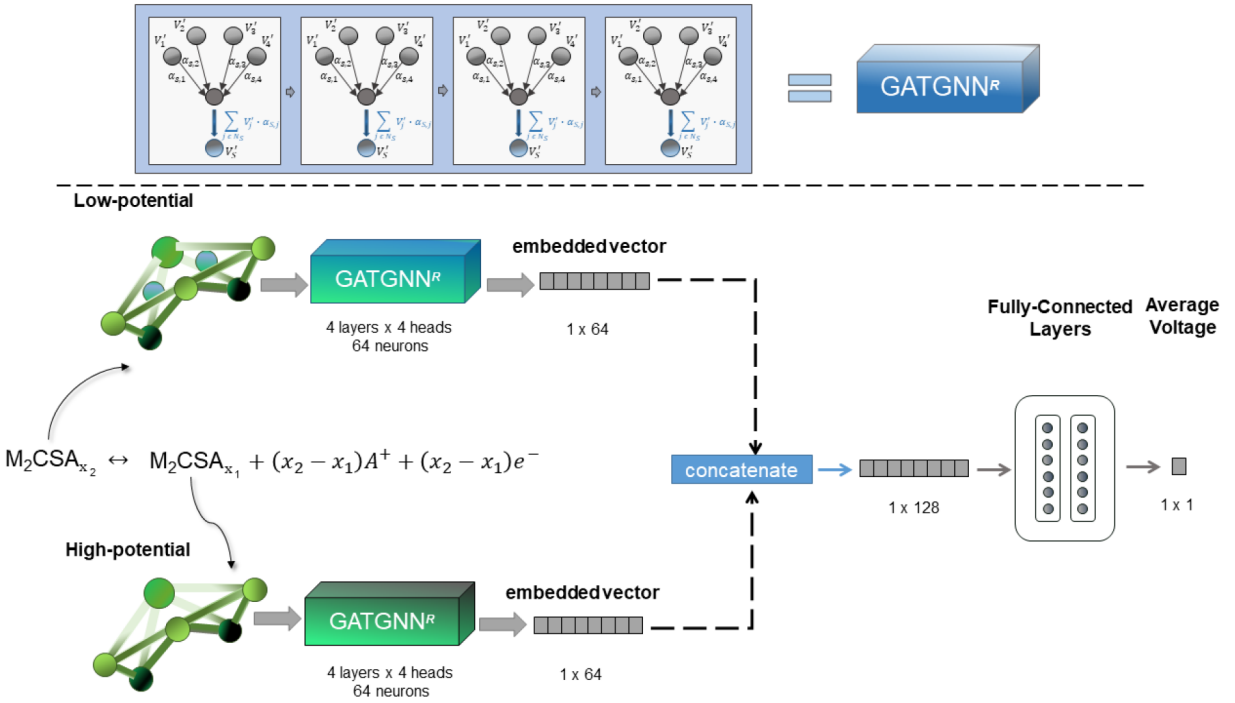


Figure 3. Architecture of our reaction-based average voltage model. The top panel shows the underlying GATGNN modules used in the work.

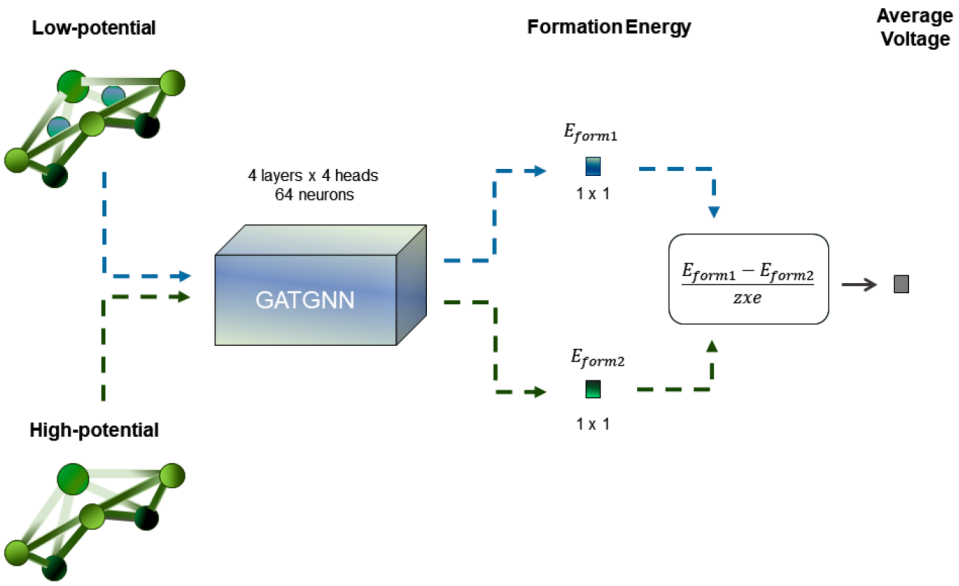


Figure 4. Architecture of E_{form}-based average voltage model.

Table 1. MAE Results for the 10-Fold Cross-Validation Study Comparing Our Proposed Reaction-Based Model to the Other Method's Composition-Based Model (Average MAE and Std Listed in the Last Column)

model	1	2	3	4	5	6	7	8	9	10	avg and std
reaction-based	0.34	0.34	0.35	0.32	0.35	0.36	0.39	0.33	0.34	0.31	0.34 ± 0.02
composition-based[16]	0.42	0.39	0.38	0.38	0.43	0.39	0.38	0.39	0.39	0.37	0.39 ± 0.02

$$\mathrm{MAE} = \frac{1}{N}\sum_{i=1}^{N} |V_i^{\mathrm{DFT}} - V_i^{\mathrm{ML}}| \tag{7}$$

where V_i^{DFT} represents the voltage computed from DFT, V_i^{ML} the machine learning predicted voltage, i a given battery sample, and N the total number of samples in a data set. Table 1 displays the cross-validation results obtained from our first experiment and the reported results from the composition-based experiments.[16] Whereas the data splits may be different, the folds were obtained from the same electrodes data set. Based on the listed MAE values, our proposed reaction-based model outperforms the composition-based model by more

https://doi.org/10.1021/acsami.2c00029
ACS Appl. Mater. Interfaces 2022, 14, 26587−26594

than 13.6%. Our reaction-based model achieved an overall average MAE of 0.34 with 0.02 standard deviation whereas the composition-based mode did 0.39 with the variance. From the results, we can conclude that our proposed method considerably benefits from the additional structural information and the learned reaction of the two electrodes.

In our second experiment, the testing set consisted of 329 battery systems comprised of 177 Li, 35 Ca, 3 Cs, 7 Rb, 11 K, 10 Y, 20 Na, 11 Al, 29 Zn, and 26 Mg battery systems. To prevent data leakage, we made sure to remove all electrodes from the 60000 size data set used to pretrain the E_{form}-based model. The same testing set aforementioned is used for evaluating our E_{form}-based approach. Figure 5 reports the

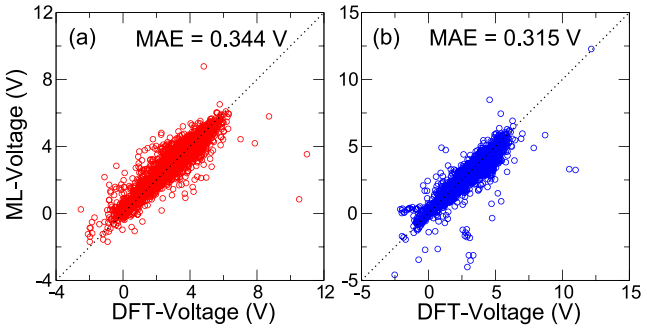


Figure 5. Parity plot comparing the DFT (DFT-Voltage) to the ML (ML-Voltage) voltage for (a) the reaction-based and (b) E_{form}-based models.

corresponding MAE and the parity plot comparing the DFT and ML voltage for both the reaction-based and the E_{form}-based models from the second experiment. As it can be seen, both reaction-based (in red) and E_{form}-based (in blue) models achieve good performance with a MAE of 0.34 and 0.31. Even though the models used in our two different proposed methods are trained by using significantly different training sets, both in terms of size (60000 vs 4402) and data (electrodes vs nonelectrodes), it is noteworthy to observe that they both achieve significantly lower MAE performance than the initially proposed composition-based model.[16] The holdout test experiment resulted in a MAE value of 0.315 for the E_{form}-based model and 0.344 for reaction-based model. While the MAE obtained from E_{form}-based model is about 8% lower than the reaction-based approach, this improvement comes at the cost of using more than 13 times more training samples (see Figure 5). Nonetheless, the E_{form}-based model does not learn directly from the battery electrodes. Therefore, we suggest that neither method is superior to the other, but instead they are alternatives. In situations where there is a lack of battery electrodes, the E_{form}-based method can be a good way to study the voltages of intercalation reactions.

To examine the transferability of our models, we evaluated our model's performance at predicting the voltage in Na- and K-ion-based electrodes. For this, we replaced the Li ions in known Li electrodes structures from the MP database with Na and K ions. For such structures, geometry optimization was not performed to calculate the voltage from our ML models. This approach allows us to examine the effect of geometry optimization on the performance of our models while predicting the voltage. Our approach is motivated by benchmark ML models used for molecules, which have achieved chemical accuracy with empirically obtained 3D coordinates from SMILES strings of molecules.[35]

From the MP database, we extracted only the DFT voltage of those Na- and K-ion-based electrodes, which share the same structure (symmetry, space groups) as the corresponding Li-based electrodes. As listed in Table 2, the Na ion shares a

Table 2. Number of Na and K Battery Frameworks That Share the Same Structures of Li Frameworks (Test Set Size) and the MAE Values in V for Predicting Voltage of the Battery Systems, Where Li Ions Were Substituted by Those Alkali Ions

		MAE (GATGNN)	
alkali ion	test set size	reaction-based	E_{form}-based
Na	127	0.56	0.47
K	32	0.87	0.70

relatively large amount of Li electrode structures (127), while K-ion electrodes have only 32. We replaced Li in those Li-based electrodes with Na and K ions and calculated the voltage using our ML models. We compared such voltage for Na- and K-based electrodes with the corresponding DFT voltage taken from the MP database (see Figure 6).

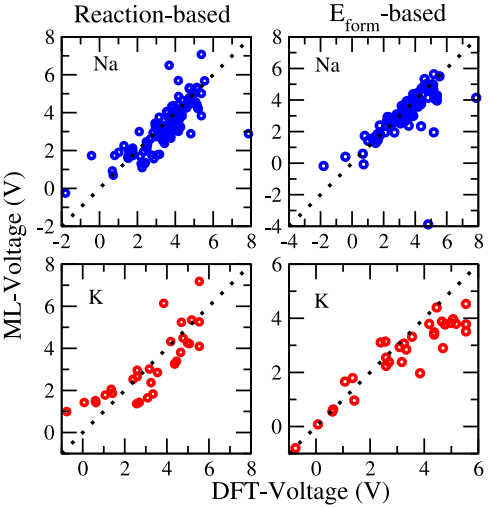


Figure 6. Performance of reaction-based and E_{form}-based models to predict the voltage of Na- and K-ion-based electrodes.

From these examinations, we obtained a somewhat large MAE of 0.56 V from the reaction-based model. For the K-ion battery electrodes, we obtained moderately large errors from both reaction-based and formation-energy-based models. We found that formation-energy-based models are more accurate at predicting voltage for a K-ion battery with a MAE of 0.70 V. The atomic radius difference between K and Li atoms is larger than that between Na and Li atoms. Therefore, the electrodes where Li was replaced by K exhibit more significant error compared to that from Na counterparts, when we predict the voltage without optimizing the atomic positions and the geometry as above. Although first-principles calculations always require geometry and cell parameter optimization before calculating voltage, our proposed E_{form}-based approach is capable of predicting the voltage of Na battery frameworks without the optimization procedure.

https://doi.org/10.1021/acsami.2c00029
ACS Appl. Mater. Interfaces 2022, 14, 26587−26594

Table 3. Comparison of the Voltages Calculated Using DFT (V_{DFT}) and That Predicted by the E_{form}-Based Model (V_{GATGNN})[a]

formula-high	formula-low	V_{DFT}	V_{GATGNN}	V_{DNN}	$\lvert\Delta V_{GATGNN}\rvert$	$\lvert\Delta V_{DNN}\rvert$
$NaMn_3OF_8$	$Na4Mn_3OF_8$	3.15	3.42	3.20	0.27	0.05
$NaCuF_4$	Na_2CuF_4	4.31	4.56	4.01	0.25	0.30
Na_2CrO_4	Na_4CrO_4	1.46	1.27	3.07	0.19	1.61
$TiCrO_4$	Na_2TiCrO_4	1.54	2.28	3.31	0.74	1.77
$Mn_3P_6WO_{24}$	$Na_6Mn_3P_6WO_{24}$	3.47	3.54	3.50	0.07	0.03
$CrWO_6$	Na_2CrWO_6	3.75	3.41	3.10	0.34	0.65
$NaVTe(WO_6)_2$	$Na_4VTe(WO_6)_2$	2.90	2.78	3.20	0.12	0.30
$NaV2O_4$	Na_2VO_2	−0.35	−0.21	2.85	0.14	3.20
$CrPO_4F$	$Na_3Cr_2P_2(O_4F)_2$	3.26	3.13	3.75	0.13	0.49
NbO_2F	$NaNbO_2F$	1.36	1.81	3.31	0.45	1.95
$NaTiV_3O_{10}$	$Na4TiV_3O_{10}$	2.47	2.32	3.00	0.16	0.53
Mn_2CrO_6	$Na_3Mn_2CrO_6$	2.73	2.73	2.95	0.00	0.22
$NaMn(PO_4)_2$	$Na3Mn(PO_4)_2$	4.32	4.47	3.28	0.15	1.04
VF_5	Na_2VF_5	4.01	4.32	3.61	0.31	0.40
$NaNbTe_2WO_{12}$	$Na_4NbTe_2WO_{12}$	3.10	3.14	3.25	0.04	0.15
$CrWO_6$	Na_2CrWO_6	3.75	3.41	3.10	0.34	0.65
$NaSbF_6$	Na_3SbF_6	2.07	3.05	3.79	0.98	1.72
$NaNbF_6$	Na_3NbF_6	3.31	1.85	2.26	1.46	1.05
				MAE	0.34	0.89

[a]The DFT calculations were performed by Moses et al.[16] after optimizing the electrodes where Li ions were replaced by Na. Our V_{GATGNN} values were predicted without relaxing the structures. We also provide the voltages for respective electrodes predicted by the composition-based DNN model (V_{DNN}).[16] $\lvert\Delta V_{GATGNN}\rvert$ is the absolute difference between V_{DFT} and V_{GATGNN}, while $\lvert\Delta V_{DNN}\rvert$ is the absolute difference between V_{DFT} and V_{DNN}. The voltage values and absolute voltage differences were calculated in V.

To further study the accuracy in predicting the voltage of the battery frameworks where Na substituted Li, we compared the predicted voltage from our E_{form}-based model (V_{GATGNN}) to the DFT calculations (V_{DFT}) reported by Moses et al.,[16] as shown in Table 3. In works done by Moses et al.,[16] the authors screened the MP database to find Na-ion battery electrodes with high voltage and minimal volume change during the charging and discharging using only the composition of electrode materials. We find that 18 battery frameworks of their data set are in our list of new Na electrodes and listed in Table 2. We also compared our predictions with the voltage predicted with simple deep neural network from their research referenced as V_{DNN}.[16] In Table 3, $\lvert\Delta V_{GATGNN}\rvert$ is the absolute difference between V_{DFT} and V_{GATGNN}, while $\lvert\Delta V_{DNN}\rvert$ is that between V_{DFT} and V_{DNN}. It is clear that the error from our model ($\lvert\Delta V_{GATGNN}\rvert$, 0.34 V) is considerably smaller than $\lvert\Delta V_{DNN}\rvert$ (0.89 V). We also found that the composition-based model (V_{DNN}) predicts a large positive voltage of 2.85 V for the NaV_2O_4 and Na_2VO_2 electrode pair, whereas our model was able to correctly predict the negative voltage of −0.21 V, which is close to the DFT predicted value of −0.35 V. This shows that our approach can more accurately represent the energetics of electrode materials compared to traditional composition-only-based models.

Another data set that we considered was published by Ong et al.[36] to study the voltages of Li- and Na-ion intercalation materials listed in Table 4. In ref 36, DFT calculations of NaMPO4 (M = Fe, Mn, Co, and Ni) materials with Olivine and Maricite structures were performed. Those two classes of materials have *Pnma* space group symmetry. Even though both have similar structures, Na and M atoms switch their sites in the two material types (i.e., in maricite structures, Na is at M's site and M is at Na's site of olivine structure). However, the DFT calculations in Table. 4 show that both structures exhibit the similar theoretical voltages for a given compound. It is clear

Table 4. Comparison of the Voltages Calculated by Ong et al. (V_{DFT})[36] and That Predicted by the E_{form}-Based Model (V_{GATGNN})[a]

material	structure	V_{DFT}	V_{GATGNN}	$\lvert\Delta V_{GATGNN}\rvert$
$NaFePO_4$	olivine	3.08	3.36	0.28
$NaMnPO_4$	olivine	3.59	3.85	0.26
$NaCoPO_4$	olivine	4.19	3.71	0.48
$NaFePO_4$	maricite	3.13	3.04	0.09
$NaMnPO_4$	maricite	3.48	3.71	0.23
$NaCoPO_4$	maricite	4.09	4.05	0.04
			MAE	0.23

[a]$\lvert\Delta V_{GATGNN}\rvert$ is the absolute difference between V_{DFT} and V_{GATGNN}. Here, the DFT calculations were performed after optimizing the electrodes where Li ions were replaced by Na. Our V_{GATGNN} values were predicted without relaxing the structures. The voltage values and absolute voltage differences were calculated in V.

that the E_{form}-based model predicted voltages provide a good accuracy (MAE = 0.23 V).

Finally, we predicted the voltages of new Na electrode materials with the $(PO_4)^{3-}$ polyanion group. These are an interesting class of materials as inductive effects of $(PO4)^{3-}$, $(P_2O_7)^{4-}$, and $(SO4)^{2-}$ polyanions offer high operating voltage. As an example, DFT calculations of $Na_xMnM'(PO_4)_3$ (M′ = Cr, Ti, and Zr) show that an average voltage around 4 V can be obtained. $Na_3V_2(PO_4)_2F_3$ fluorophosphate has a theoretical voltage of 3.9 V. Thus, $(PO_4)^{3-}$-based electrodes are popular as high-density cathode materials for Na-ion batteries.[37] $Na_3(MPO_4)_2F_3$ (M = Al, V, Fe, Cr, and Ga) and $AVPO_4F$ (A = Na and Li) have been widely studied as the cathode materials for Li and Na batteries.[38,39] Table 5 contains predicted voltages for 10 fluorophosphate materials, which are not included in the MP database. The $NaTiPO_4F$ and $TiPO_4F$ pair of electrodes provides a considerably lower voltage than that of the same family of electrodes with the other transition metal atoms. Ti-based compounds have been widely

https://doi.org/10.1021/acsami.2c00029
ACS Appl. Mater. Interfaces 2022, 14, 26587−26594

Table 5. Voltages Predicted by E_{form}-Based Model for 10 Fluorophosphate-Based Battery Frameworks, Which Are Not Included in the MP Database for Na Batteries (Voltage Values Calculated in V)

formula-low	formula-high	V_{GATGNN}
$NaTiPO_4F$	$TiPO_4F$	1.32
$NaCrPO_4F$	$CrPO_4F$	3.48
$NaFePO_4F$	$FePO_4F$	3.22
$NaCuPO_4F$	$CuPO_4F$	5.52
$NaMnPO_4F$	$MnPO_4F$	4.48
Na_2CoPO_4F	$NaCoPO_4F$	3.48
Na_2MnPO_4F	$NaMnPO_4F$	3.42
$Na_3Co_2(PO_4F)_2$	$NaCoPO_4F$	3.88
$Na_3Ni_2(PO_4F)_2$	$NaNiPO_4F$	3.51
$Na_3Cr_2(PO_4F)_2$	$CrPO_4F$	3.13

investigated as anode materials for Na-battery electrodes due to low operating voltages and stability.[40] It is reported that $NaTiO_2$ (1.37 V) electrode has a low theoretical voltage, even though Ni- and Co-based counterparts have voltages higher than 3.3 V.[36] The battery framework with the $NaCuPO_4F$ and $CuPO_4F$ pair of electrodes shows the highest voltage, which is 5.52 V, in Table 5. All the other battery systems in Table 5 exhibit theoretical voltages between 3.1 and 4.5 V.

4. CONCLUSION

In summary, we developed two attention-based graph neural networks that combine the chemical compositions with the spatial information to predict the voltage of the battery electrode materials. The first method predicts the voltage by considering the chemical reaction between a high-potential electrode and the metal ions to form a low-potential electrode. The second model predicts the E_{form} of individual electrodes before being used to compute the voltage. Results were compared with the latest composition based model (Moses et al.[16]) from the literature for predicting voltage. Our structure-based models are much more accurate than this benchmark work. Our E_{form}-based model consistently provides lower MAE compared to that from reaction-based model. And also, we show that relative to known models in the literature, our E_{form}-based model demonstrates high transferability of performance when applied to Na electrodes. Furthermore, we predicted the average voltages of 10 fluorophosphate-based battery frameworks, which are not included in the MP database for Na batteries. Those fluorophosphates have the Na_xMPO_4F (M = Ti, Cr, Fe, Cu, Mn, Co, and Ni) general chemical formula. We could show that it can expect average voltages greater than 3.1 V from those Na battery frameworks except from the $NaTiPO_4F$ and $TiPO_4F$ pair of electrodes, where it exhibits an average voltage of 1.32 V.

AUTHOR INFORMATION

Corresponding Authors

Rajendra P. Joshi − *Pacific Northwest National Laboratory, Richland, Washington 99352, United States; TQuT Inc., Rockford, Michigan 49341, United States;* Email: rajendra.joshi@pnnl.gov

Jianjun Hu − *Department of Computer Science and Engineering, University of South Carolina, Columbia, South Carolina 29201, United States;* orcid.org/0000-0002-8725-6660; Email: jianjunh@cse.sc.edu

Authors

Steph-Yves Louis − *Department of Computer Science and Engineering, University of South Carolina, Columbia, South Carolina 29201, United States*

Edirisuriya M. Dilanga Siriwardane − *Department of Computer Science and Engineering, University of South Carolina, Columbia, South Carolina 29201, United States; Department of Physics, University of Colombo, Colombo 0300, Sri Lanka;* orcid.org/0000-0001-8960-5273

Sadman Sadeed Omee − *Department of Computer Science and Engineering, University of South Carolina, Columbia, South Carolina 29201, United States;* orcid.org/0000-0002-1016-8072

Neeraj Kumar − *Pacific Northwest National Laboratory, Richland, Washington 99352, United States;* orcid.org/0000-0001-6713-2129

Complete contact information is available at: https://pubs.acs.org/10.1021/acsami.2c00029

Author Contributions

S.L. and E.S. contributed equally to this work. Conceptualization, R.P.J., E.S., and J.H.; methodology, S.L., E.S., and J.H.; software, S.L. and E.S.; validation, S.L., E.S., and R.P.J.; investigation, S.L., E.S., and J.H.; resources, J.H. and N.K.; data curation, R.P.J. and E.S.; writing−original draft preparation, E.S., S.L., R.P.J., S.S.O., and J.H.; writing−review and editing, N.K. and J.H.; supervision, J.H. and R.P.J.; funding acquisition, J.H. and N.K.

Notes

The authors declare no competing financial interest.

Codes for this work can be found at https://github.com/superlouis/GATGNN-VOLTAGE.

ACKNOWLEDGMENTS

Research reported in this work was supported in part by NSF under Grants 1940099 and 1905775. This work was also partially supported by the Department of Energy under Grant DE-SC0020272. R.P.J. and N.K. were supported by Laboratory Directed Research and Development Program and Mathematics for Artificial Reasoning for Scientific Discovery investment at the Pacific Northwest National Laboratory, a multiprogram national laboratory operated by Battelle for DOE under Contract DE-AC06-76RLO. The views, perspective, and content do not necessarily represent the official views of NSF or DE.

REFERENCES

(1) Li, M.; Lu, J.; Chen, Z.; Amine, K. 30 Years of Lithium-Ion Batteries. *Adv. Mater.* **2018**, *30*, 1800561.

(2) Guo, H.; Wang, Q.; Stuke, A.; Urban, A.; Artrith, N. Accelerated Atomistic Modeling of Solid-State Battery Materials With Machine Learning. *Front. Energy Res.* **2021**, *9*, 265.

(3) Deng, J.; Bae, C.; Denlinger, A.; Miller, T. Electric Vehicles Batteries: Requirements and Challenges. *Joule* **2020**, *4*, 511−515.

(4) Goel, S.; Sharma, R.; Rathore, A. K. A Review on Barrier and Challenges of Electric Vehicle in India and Vehicle to Grid Optimisation. *Transp. Eng.* **2021**, *4*, 100057.

(5) Jain, A.; Shin, Y.; Persson, K. A. Computational Predictions of Energy Materials Using Density Functional Theory. *Nat. Rev. Mater.* **2016**, *1*, 1−13.

(6) Gandomi, Y. A.; Aaron, D. S.; Houser, J. R.; Daugherty, M. C.; Clement, J. T.; Pezeshki, A. M.; Ertugrul, T. Y.; Moseley, D. P.; Mench, M. M. Critical Review—Experimental Diagnostics and

Material Characterization Techniques Used on Redox Flow Batteries. *J. Electrochem. Soc.* **2018**, *165*, A970−A1010.
(7) Natarajan, A. R.; Van der Ven, A. Machine-Learning the Configurational Energy of Multicomponent Crystalline Solids. *npj Comput. Mater.* **2018**, *4*, 1−7.
(8) Houchins, G.; Viswanathan, V. An Accurate Machine-learning Calculator for Optimization of Li-ion Battery Cathodes. *J. Chem. Phys.* **2020**, *153*, 054124.
(9) Allam, O.; Kuramshin, R.; Stoichev, Z.; Cho, B.; Lee, S.; Jang, S. Molecular Structure−Redox Potential Relationship for Organic Electrode Materials: Density Functional Theory−Machine Learning Approach. *Mater. Today Energy* **2020**, *17*, 100482.
(10) Artrith, N.; Urban, A.; Ceder, G. Constructing First-Principles Phase Diagrams of Amorphous Li X Si Using Machine-Learning-Assisted Sampling with an Evolutionary Algorithm. *J. Chem. Phys.* **2018**, *148*, 241711.
(11) Shandiz, M. A.; Gauvin, R. Application of Machine Learning Methods for The Prediction of Crystal System of Cathode Materials in Lithium-Ion Batteries. *Comput. Mater. Sci.* **2016**, *117*, 270−278.
(12) Cubuk, E. D.; Sendek, A. D.; Reed, E. J. Screening Billions of Candidates for Solid Lithium-Ion Conductors: A Transfer Learning Approach for Small Data. *J. Chem. Phys.* **2019**, *150*, 214701.
(13) Cunha, R. P.; Lombardo, T.; Primo, E. N.; Franco, A. A. Artificial Intelligence Investigation of NMC Cathode Manufacturing Parameters Interdependencies. *Batter. Supercaps* **2020**, *3*, 60−67.
(14) Dixit, M. B.; Verma, A.; Zaman, W.; Zhong, X.; Kenesei, P.; Park, J. S.; Almer, J.; Mukherjee, P. P.; Hatzell, K. B. Synchrotron Imaging of Pore Formation in Li Metal Solid-State Batteries Aided by Machine Learning. *ACS Appl. Energy Mater.* **2020**, *3*, 9534−9542.
(15) Joshi, R. P.; Eickholt, J.; Li, L.; Fornari, M.; Barone, V.; Peralta, J. E. Machine Learning the Voltage of Electrode Materials in Metal-Ion Batteries. *ACS Appl. Mater. Interfaces* **2019**, *11*, 18494−18503.
(16) Moses, I. A.; Joshi, R. P.; Ozdemir, B.; Kumar, N.; Eickholt, J.; Barone, V. Machine Learning Screening of Metal-Ion Battery Electrode Materials. *ACS Appl. Mater. Interfaces* **2021**, *13*, 53355−53362.
(17) Chen, C.; Ye, W.; Zuo, Y.; Zheng, C.; Ong, S. P. Graph Networks as a Universal Machine Learning Framework for Molecules and Crystals. *Chem. Mater.* **2019**, *31*, 3564−3572.
(18) Xie, T.; Grossman, J. C. Crystal Graph Convolutional Neural Networks for an Accurate and Interpretable Prediction of Material Properties. *Phys. Rev. Lett.* **2018**, *120*, 145301.
(19) Banjade, H. R.; Hauri, S.; Zhang, S.; Ricci, F.; Gong, W.; Hautier, G.; Vucetic, S.; Yan, Q. Structure Motif − − Centric Learning Framework for Inorganic Crystalline Systems. *Sci. Adv.* **2021**, *7*, No. eabf1754.
(20) Ong, S. P.; Richards, W. D.; Jain, A.; Hautier, G.; Kocher, M.; Cholia, S.; Gunter, D.; Chevrier, V. L.; Persson, K. A.; Ceder, G. Python Materials Genomics (Pymatgen): A Robust, Open-Source Python Library for Materials Analysis. *Comput. Mater. Sci.* **2013**, *68*, 314−319.
(21) Zhou, F.; Cococcioni, M.; Marianetti, C. A.; Morgan, D.; Ceder, G. First-Principles Prediction of Redox Potentials in Transition-Metal Compounds With LDA+U. *Phys. Rev. B* **2004**, *70*, 235121.
(22) Jain, A.; Ong, S. P.; Hautier, G.; Chen, W.; Richards, W. D.; Dacek, S.; Cholia, S.; Gunter, D.; Skinner, D.; Ceder, G.; Persson, K. A. The Materials Project: A Materials Genome Approach to Accelerating Materials Innovation. *APL Mater.* **2013**, *1*, 011002.
(23) Wu, Z.; Pan, S.; Chen, F.; Long, G.; Zhang, C.; Yu, P. S. A Comprehensive Survey on Graph Neural Networks. *EEE Trans. Neural Netw. Learn. Syst.* **2021**, *32*, 4−24.
(24) Louis, S.-Y.; Nasiri, A.; Rolland, F. J.; Mitro, C.; Hu, J. NODE-SELECT: A Graph Neural Network Based On A Selective Propagation Technique. *arXiv preprint arXiv:2102.08588*, **2021**.
(25) Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A. N.; Kaiser, L.; Polosukhin, I. Attention is All You Need. *arXiv preprint arXiv:1706.03762*, **2017**.
(26) Veličković, P.; Cucurull, G.; Casanova, A.; Romero, A.; Lio, P.; Bengio, Y. Graph Attention Networks. *arXiv preprint arXiv:1710.10903*, **2017**.
(27) Louis, S.-Y.; Zhao, Y.; Nasiri, A.; Wang, X.; Song, Y.; Liu, F.; Hu, J. Graph Convolutional Neural Networks with Global Attention for Improved Materials Property Prediction. *Phys. Chem.* **2020**, *22*, 18141−18148.
(28) Omprakash, P.; Manikandan, B.; Sandeep, A.; Shrivastava, R.; Viswesh, P.; Panemangalore, D. B. Graph Representational Learning for Bandgap Prediction in Varied Perovskite Crystals. *Comput. Mater. Sci.* **2021**, *196*, 110530.
(29) Rosen, A. S.; Iyer, S. M.; Ray, D.; Yao, Z.; Aspuru-Guzik, A.; Gagliardi, L.; Notestein, J. M.; Snurr, R. Q. Machine Learning the Quantum-Chemical Properties of Metal−Organic Frameworks for Accelerated Materials Discovery. *Matter* **2021**, *4*, 1578−1597.
(30) Paszke, A.; Gross, S.; Chintala, S.; Chanan, G.; Yang, E.; DeVito, Z.; Lin, Z.; Desmaison, A.; Antiga, L.; Lerer, A. Automatic Differentiation in Pytorch. *Neural Inf. Process.* **2017**.
(31) Fey, M.; Lenssen, J. E. Fast Graph Representation Learning with PyTorch Geometric. *arXiv preprint arXiv:1903.02428*, **2019**.
(32) Girshick, R. Fast r-cnn. Proceedings of the IEEE International Conference on Computer Vision, 2015; pp 1440−1448.
(33) Gallarati, S.; Fabregat, R.; Laplaza, R.; Bhattacharjee, S.; Wodrich, M. D.; Corminboeuf, C. Reaction-based Machine Learning Representations for Predicting the Enantioselectivity of Organo-catalysts. *Chem. Sci.* **2021**, *12*, 6879−6889.
(34) Coley, C. W.; Barzilay, R.; Jaakkola, T. S.; Green, W. H.; Jensen, K. F. Prediction of Organic Reaction Outcomes Using Machine Learning. *ACS Cent. Sci.* **2017**, *3*, 434−443.
(35) Hirohara, M.; Saito, Y.; Koda, Y.; Sato, K.; Sakakibara, Y. Convolutional Neural Network Based on SMILES Representation of Compounds for Detecting Chemical Motif. *BMC Bioinform* **2018**, *19*, 83−94.
(36) Ong, S. P.; Chevrier, V. L.; Hautier, G.; Jain, A.; Moore, C.; Kim, S.; Ma, X.; Ceder, G. Voltage, Stability and Diffusion Barrier Differences Between Sodium-Ion and Lithium-Ion Intercalation Materials. *Energy Environ. Sci.* **2011**, *4*, 3680−3688.
(37) Wang, J.; Wang, Y.; Seo, D.-H.; Shi, T.; Chen, S.; Tian, Y.; Kim, H.; Ceder, G. A High-Energy NASICON-Type Cathode Material for Na-Ion Batteries. *Adv. Energy Mater.* **2020**, *10*, 1903968.
(38) Zhao, J.; Mu, L.; Qi, Y.; Hu, Y.-S.; Liu, H.; Dai, S. A Phase − Transfer Assisted Solvo − Thermal Strategy for Low − Temperature Synthesis of $Na_3(VO_{1x}PO_4)_2F_{1+2x}$ Cathodes for Sodium − Ion Batteries. *Chem. Commun.* **2015**, *51*, 7160−7163.
(39) Barker, J.; Saidi, M. Y.; Swoyer, J. L. A Sodium-Ion Cell Based on the Fluorophosphate Compound NaVPO[sub 4]F. *Electrochem. Solid-State Lett.* **2003**, *6*, A1.
(40) Fang, Y.; Xiao, L.; Chen, Z.; Ai, X.; Cao, Y.; Yang, H. Recent Advances in Sodium-Ion Battery Materials. *EER* **2018**, *1*, 294−323.

https://doi.org/10.1021/acsami.2c00029
ACS Appl. Mater. Interfaces 2022, 14, 26587−26594