

Machine Learning as a Tool for Specific Capacity Prediction of Prospective Potassium Battery Electrodes

Souvik Manna[†], Diptendu Roy[†], Sandeep Das[†], and Biswarup Pathak^{†*}

[†]Department of Chemistry, Indian Institute of Technology (IIT) Indore, Simrol, Indore 453552, India

*Email: biswarup@iiti.ac.in

ABSTRACT

Application of data science and machine learning (ML) techniques in the domain of materials science has been increasing by leaps and bounds recently. With the help of ML, through input features derived from available databases we can rapidly screen materials based on our desired output. Capacity is one of the important parameters for choosing suitable electrode materials for high energy storage metal ion battery. Exploration of suitable electrode materials for metal ion batteries other than Li ion batteries (LIBs) has been deficient, though there is a need to develop alternative battery technologies with higher energy storage characteristics and environmental safety. In this work, we have considered Li, Na and K-ion electrode materials and their available battery data from Materials Project database to predict specific capacity of prospective K-ion battery electrode materials. Suitable features have been considered and developed to train the various ML algorithms. Mean Absolute Percentage Error has been considered as the performance metrics for assessment of different ML algorithms and among them, kernel ridge regression has been adopted as the most useful to predict the capacity of unknown electrode materials for K-ion

battery. Using the value of specific capacity, the number of intercalated K ions in the formula unit of the non-intercalated electrode material compounds have also been calculated. DFT calculations have also been performed to verify the results obtained through ML. Our result shows ML is an encouraging alternative to computationally demanding DFT process as it can screen electrode materials rapidly for battery.

Keywords: machine learning, batteries, data analysis, electrode materials, capacity prediction.

1. Introduction

With the increase in energy demands, harnessing energy from renewable sources has become increasingly important for sustainable development. However, renewable energy sources are intermittent in nature as they depend on factors like weather, location, efficiency and available infrastructure. Thus, efficient energy storage systems, especially for large-scale are required to store, transfer and utilize the energy produced from renewable energy sources¹. Rechargeable metal-ion batteries are used extensively to store energy in the form of chemical energy which can be converted back to electrical energy whenever required. Among all the metal-ion battery, Li-ion batteries (LIBs) are leading the energy storage devices market, especially in portable devices such as smartphones and laptops^{2,3}. Li metal-ion batteries has even opened extraordinary possibilities in automotive sector and electric vehicles market recently.^{4,5} The long cycle life, high efficiencies and high energy densities are the main reason behind the success of LIBs.^{2,6} However, for large scale energy storage, LIBs have certain shortcomings such as its relatively low energy density, and safety issue owing its high reactivity in air.^{4,6–11} Very low abundance of Li sources is also a major

concern which ultimately contributes to the high price of these batteries.^{6,12,13} These issues demand for cheap, efficient and sustainable alternatives of LIBs.

Potassium is one of the metal ions which could replace lithium in energy storage devices. K is more abundant compared to Li sources and hence reduces the production cost.¹⁴ K-ion batteries have a similar rocking chair mechanism like LIBs. K^+ having large atomic radius (1.38 Å) has a small Stokes radius in various organic electrolytes which results in higher ionic conductivity.¹⁵ Though the electrode materials for LIBs have been extensively explored, but the same is not true for K-ion batteries. However, seeking suitable electrode materials for K ion battery is experimentally challenging and even theoretically, requires high computational facilities. Majority of the electrode materials used for LIBs are still unexplored for K-ion battery due to the difficulties in experimental and computational screening of large number of electrode materials with high accuracy.^{16–18} Therefore, machine learning (ML) could be an advanced tool which can save both time and cost, and at the same time screen many electrodes with minimum computational cost. For the exploration of electrode materials by ML, verified and well-curated data is needed, which is enabled by the density functional theory (DFT) based databases like NOMAD, OQMD, AFLOW, Materials Project.^{19–25} Although use of DFT based data is not standard for every context, still it delivers sensible insights which ultimately helps in the guidance of experimental research.^{26,27} ML combined with data from various databases can be used for predicting any specific property of interest for a particular battery material.^{17,23,28–32} Application of ML in the field of material science can be found in the prediction of microscopic properties like band structure, formation energy, solar cells, batteries, density of states and catalysis.^{33–48} Kernel ridge regression (KRR) and support vector regression (SVR) has been used by Seko et al. for the prediction of thermal conductivity and cohesive energy of binary and tertiary compounds.^{49,50} ML techniques have also been used for

the prediction of different properties for their applicability as materials in photovoltaic cells and glass alloys.^{51–53} Sendek et al., used the logistic regression for proposing 12000 Li containing solids as solid-state electrolytes for LIBs by rapid screening.⁴⁵ Meredig et al. built ML model to estimate thermodynamic stability and proposed around 4500 stable novel materials.⁵⁴ ML has also been used for the improvement of DFT , prediction of thermal, electronic and mechanical properties.^{23,55–57}

Capacity is one of the important metrics for the measurements of battery performance. The longevity of a battery mainly depends on cycle life of a battery and the former directly related to the capacity of a battery. From the number of ions intercalated in an electrode material we can find out the capacity of that electrode material and in order to do so quantum mechanically we need to perform DFT calculations for each individual electrode material which is very time-consuming process. However, we can utilize the different advance machine learning model as a tool to speed up the screening of electrode materials based on capacity as target variable. As far best of our knowledge till now there are very few studies has been carried out on the capacity using ML via cycle life for a particular electrode material.^{58,59} However, there are hardly any work has done directly on capacity of different electrode materials by ML. In this study we have utilized the Li, Na, and K ion battery data for the training of ML models in order to predict the capacity of those electrode materials for the K ion battery. The capacity of different electrode materials varies rapidly, and the range of minimum capacity and maximum capacity is very high. Keeping that in mind, we have only considered the monovalent ions and not bivalent and trivalent ions for intercalation. We have also not considered the lower alkaline metal ions since the radius of those ions will increase as we go down the group. Among the metal ion batteries, LIBs have been explored extensively, however, experimentally or by DFT calculation testing all those LIBs

electrode materials for K ion batteries is a lengthy process. Therefore, after considering the Li, Na and K ion battery data for the training set we have replaced the Li and Na by K for an approx. estimation of capacity with the help of different machine learning models. Here, we have used Support vector machine (SVM), ExtraTrees regressor (EXR) and kernel ridge regression (KRR) to fit training dataset. With addition to our particular interest that is capacity, we further used the predicted capacity for the calculation of number of K ion could be intercalated in the LIBs and NaIBs electrode materials. The performance of different ML model was assessed by mean absolute percentage error (MAPE). DFT calculation for few unknown electrode materials has been performed to validate the machine learning model.

The training data for these metal ion batteries has been retrieved from the materials project database using pymatgen (Python Materials Genomics).^{24,60,61} The data set consists of 69.54% of Li, 22.4% of Na and 8.06% of K ion battery data. So, from the amount of data, there is a high chance of getting the target variable that is capacity to mimic the electrodes used for the LIBs since the contribution in the overall data from LIBs is very high compared to other two metal ion batteries. The overall known dataset is divided in training set and validation set. The training set has been used to train the ML model whereas the validation set was used to validate the performance of our machine learning models. The validation set is composed of 20% of total data and rest of the data used for training. The training set remains unique for all the ML model used and same is true for the validation set. The distribution of the training set of metal ion batteries electrode materials has been shown in Figure 1. For each electrode materials we have generated 196 unique elemental descriptors depending on chemical formula of individual electrode materials using choice-based feature vectorization.⁶² Along with these descriptors other structural parameters like lattice parameters (a, b, c), lattice angles (α , β , γ), volume of void etc. have been

also considered, so that these descriptors can represent each electrode materials uniquely. In order to specify the intercalated ion, we have also included some elemental properties of intercalated ions like ionic radius, ionization energy, heat of atomization, etc. After the generation of descriptors scaling has performed on each descriptor except on target variable using StandardScaler module of python package to bring down the all the descriptors in the same scale so that we can avoid the biasness of our data set based on the magnitude of each descriptor of electrode materials by machine.

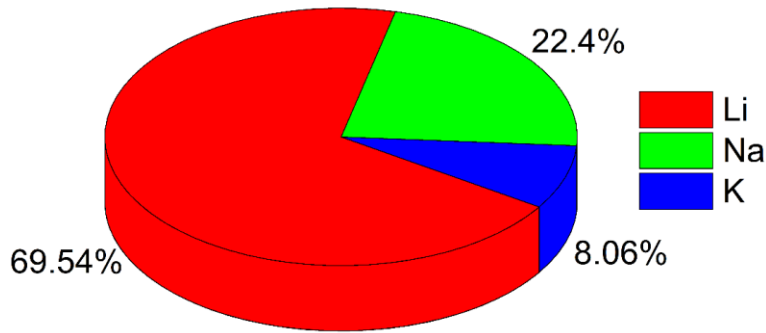


Figure 1: Distribution of different metal ions battery data used in our ML model for training.

Further, whether all features are important or not for the desired target variable, we have performed Lasso Regression. With the help of Lasso regression, we have calculated the feature importance of each individual descriptors and based on the magnitude of the feature importance we have screened the features. We have only selected those features having features greater than zero and rest of the features have been eliminated for the fitting of ML models. The mathematical expression of LASSO Regression is given as

$$\sum_{i=1}^n (y - y')^2 + \lambda m$$

where y is the actual value and y' is the value of best fitted line. By Lasso regression we have calculated the slope (m) value for each feature. λ is a constant and considering its value equal to one, the slope for each feature has been established. Those features having m value equal to zero were considered as irrelevant and those features were dropped. The lasso regression helped to shrink the feature by 64%. The selection of features based on feature importance has shown in the Figure 2.

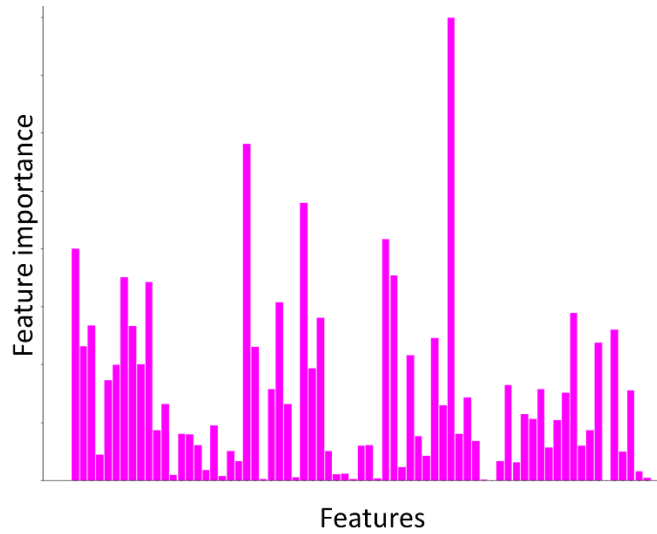


Figure 2: Selection of features by Lasso Regression.

2. Data Analysis

For finding out the correlation among the features we have also generated the heat map. The heat map has shown in Figure 3 was generated by the correlation function from the seaborn library. The features are listed in the supported information. From the correlation values of different features in the heat map (Figure 3), it has been observed almost all the features are independent of each other, some descriptors are positively correlated, and some are negatively. This result proves the choice of these descriptors will be able to represent each electrode materials uniquely. Though

there are very few elemental descriptors which are dependent of each other, we can't drop those descriptors since those descriptors represents the ions which are getting intercalated. The list of elemental features used in the heat-map (f_i) is attached with the Table S1 (Supporting Information).

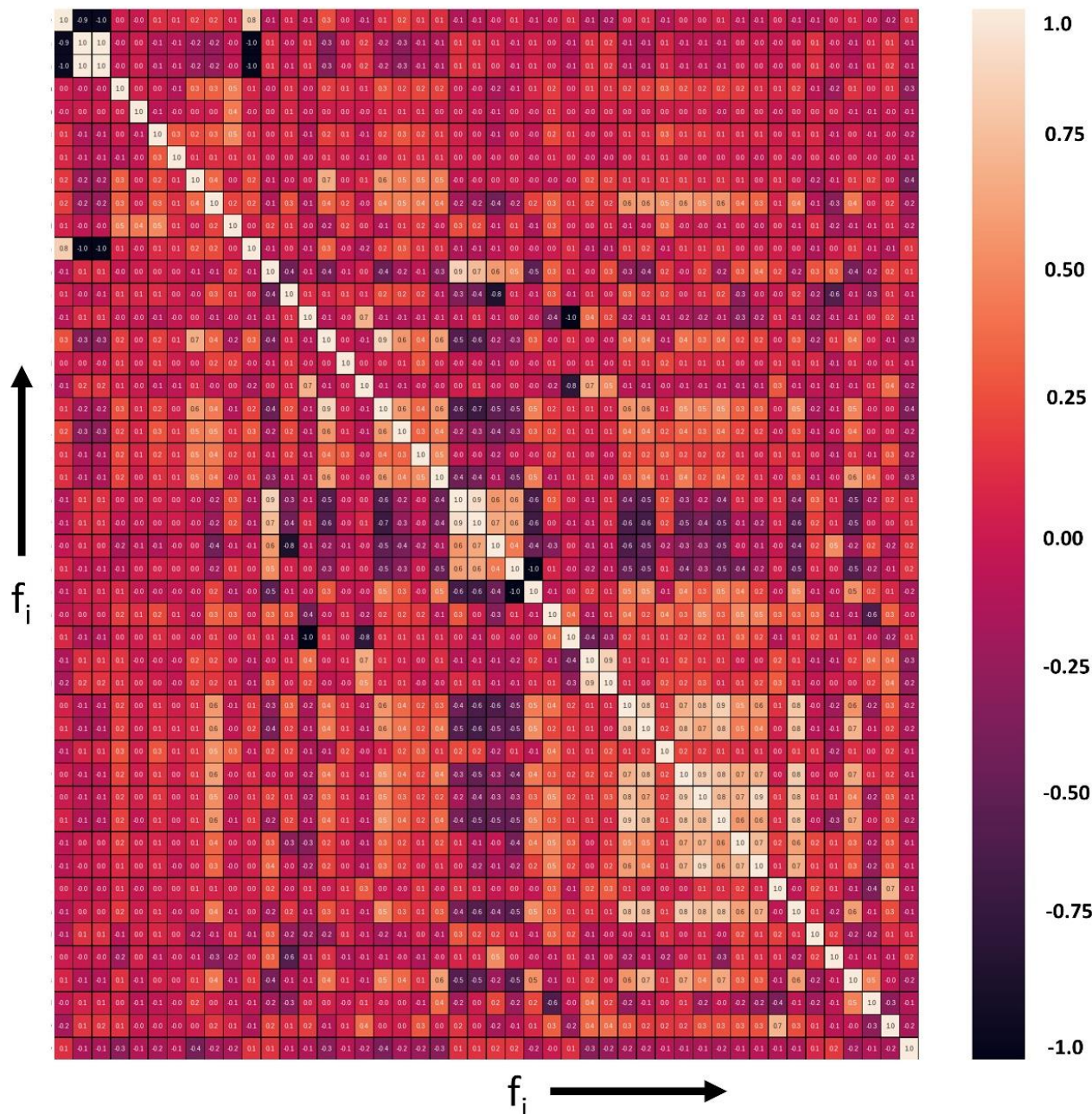


Figure 3: Heatmap showing the correlation among the considered features.

Understanding the nature of different features is highly important and, in this regard, distribution plot help us to find out the density of every feature. Here we have shown distribution plot of some selected descriptors as these descriptors are very impactful on determining the capacity, rest of the plots are attached with the supporting information (SI). Some of the descriptors has been taken to produce joint plot across the capacity of the electrode for finding out some insight and dependency of the dependent variable on the independent variable. The joint plot of the selected descriptors are taken by intuition as we were keen to know whether these descriptors plays important role from chemical point of view or not.

From the Figure 4, it is observed that molecular weight and polarizability follows almost same trend whereas specific heat values are diverse at higher magnitude. There is an indirect correlation between molecular weight and polarizability as electron density increases with the increase in atomic mass. Therefore, it is very likely to observe the similar tendency between these two parameters. We have not observed any trend in change of capacity with respect to average Pauling electronegativity. The high range of electronegativity for most of the electrode materials can be the cause for this as it overestimates the capacity for the overall data set with respect to this feature. The electronic properties also play important role for electrode materials. To understand how the capacity of electrode materials change with the change in the electronic properties of those materials, we have calculated the average valence s, d, f electrons by taking the average of the valence electrons of the constituent atoms of electrode materials and plotted the contribution of valence electrons with capacity change. From the Figure 5, it is observed that though the average capacity falling in the range of higher s electron contribution, completely opposite trend observed in case of d orbital valence electron. So, in general there are large number of electrodes having high s orbital valence electrons and low d orbital valence electrons but there is no such trend is

observed for the f orbital valence electron as the capacity range changes from low magnitude to high magnitude at constant f orbital valence electron. The reason may be most of the electrode consists of transition metals having high number of d electrons with filled s orbital whereas the f orbital remains almost empty and even though for some electrode materials there exist some f electrons since they are deeply seated might not be as effective as the s or d valence electrons.

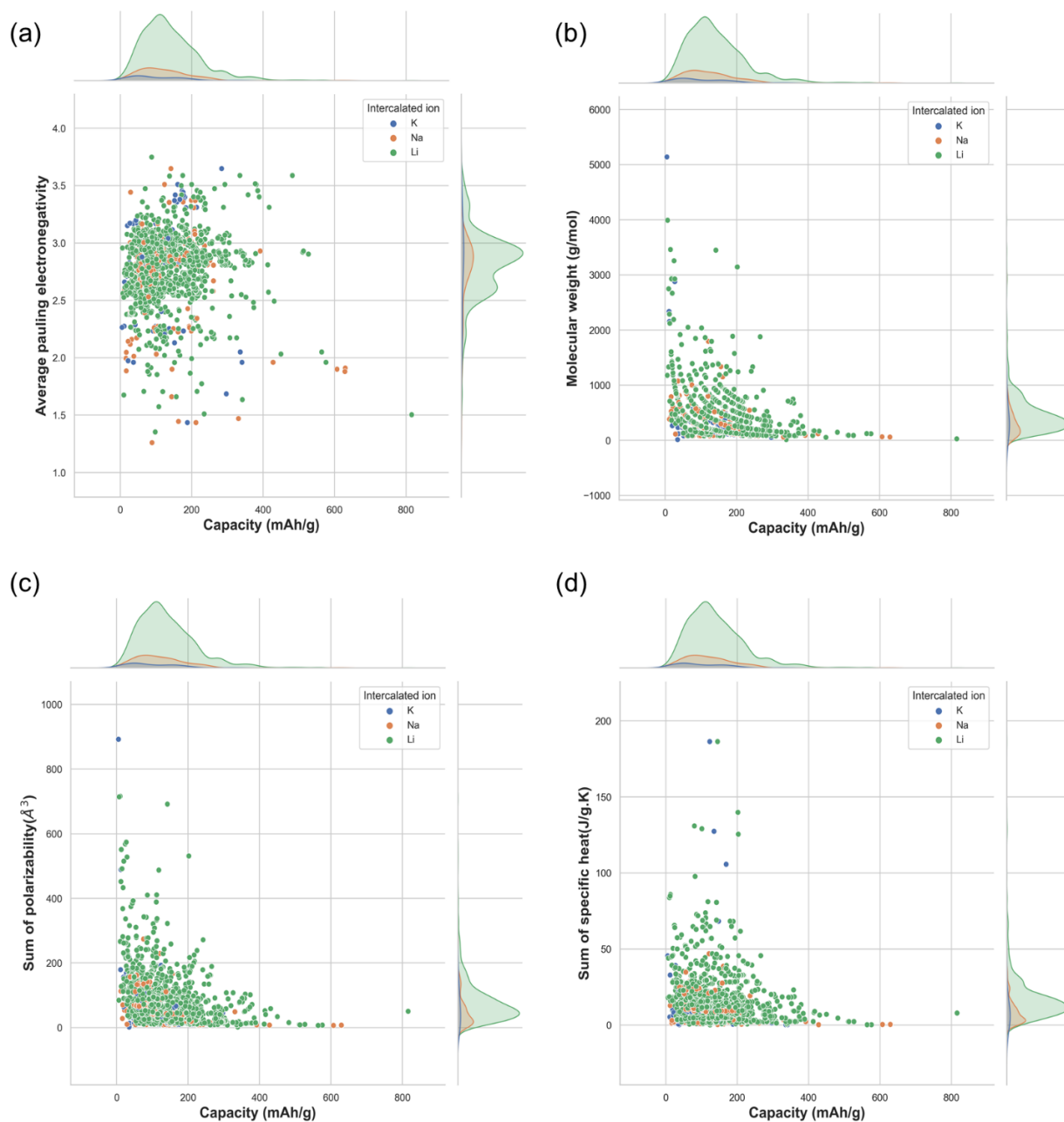


Figure 4: Joint plot for the density and distribution of capacity with respect to molecular properties.

change in capacity with (a) average Pauling electronegativity, (b) molecular weight, (c) sum of polarizability, and (d) sum of specific heat.

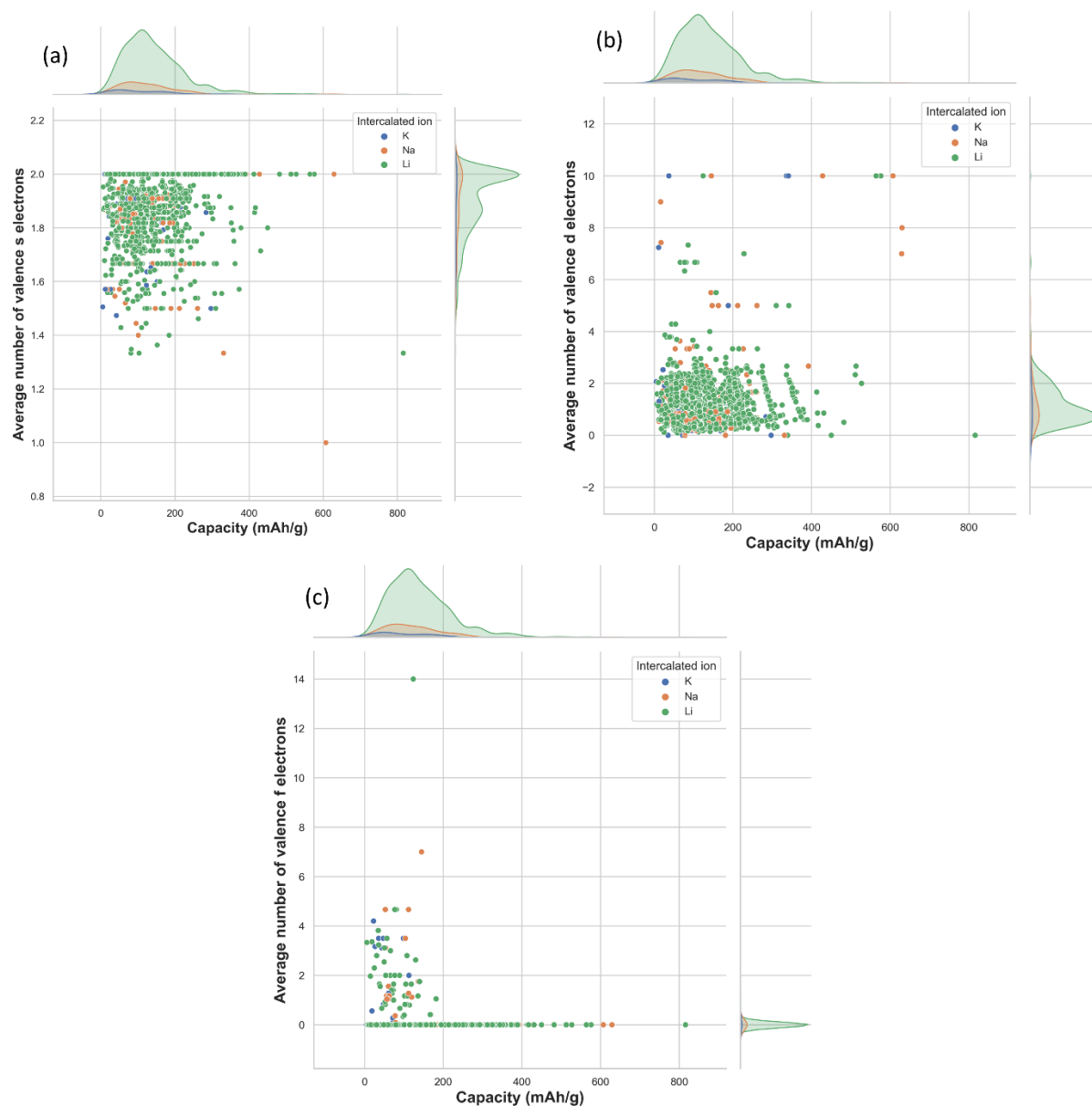


Figure 5: Joint plot for the distribution plot across electronic properties. Change of capacity with

(a) s valence electrons (b) d valence electrons (c) f valence electrons.

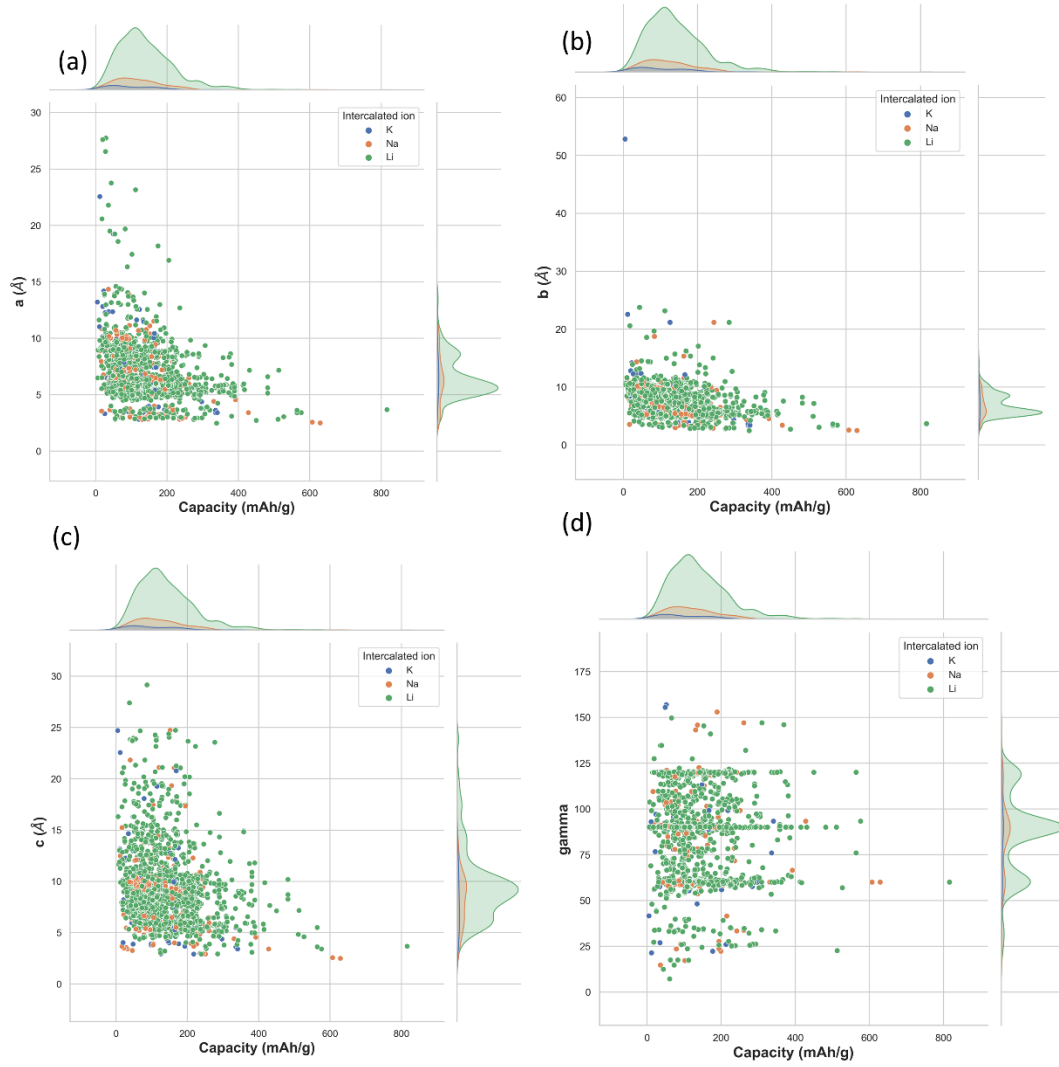


Figure 6: Distribution of capacity with respect to different lattice parameters of electrode materials.

(a) Change in capacity with lattice parameter a , (b) Change in capacity with lattice parameter b , (c) Change in capacity with lattice parameter c , (d) Change in capacity with lattice angle γ .

In the Figure 6, we have shown how the distribution of capacity changes with the different lattice parameters (a , b , c) and lattice angle (γ) of electrode materials. From the plot it is observed, the effect of lattice parameter a and c is very prominent compared to the lattice parameter b . This may be because for most of the electrode materials in one direction remain fixed like b here and

the other two direction a and c has changed for the intercalation of metal ions. Thus, we can say most of the electrode varies in two directions keeping the length in one direction constant. In case of gamma, though it changes rapidly but the capacity changes almost in the same range for low and high value of gamma. We have also plotted the box plot between Gravimetric capacity and the ionic radius of intercalated ions (Figure 7) to get a visualization of average capacity range of Li, Na, and K ion battery. The mid-line in the box plot is the median, lower line outside the box is the minimum range and the upper line outside the box is the maximum range of our property of interest. Therefore, in the above plot we can see the average capacity for Li ion battery is higher followed by Na and K ion batteries and it is expected since we know with decrease in ionic radius the number of ions intercalated within an electrode material will be higher and so the capacity of that electrode for that particular ion.

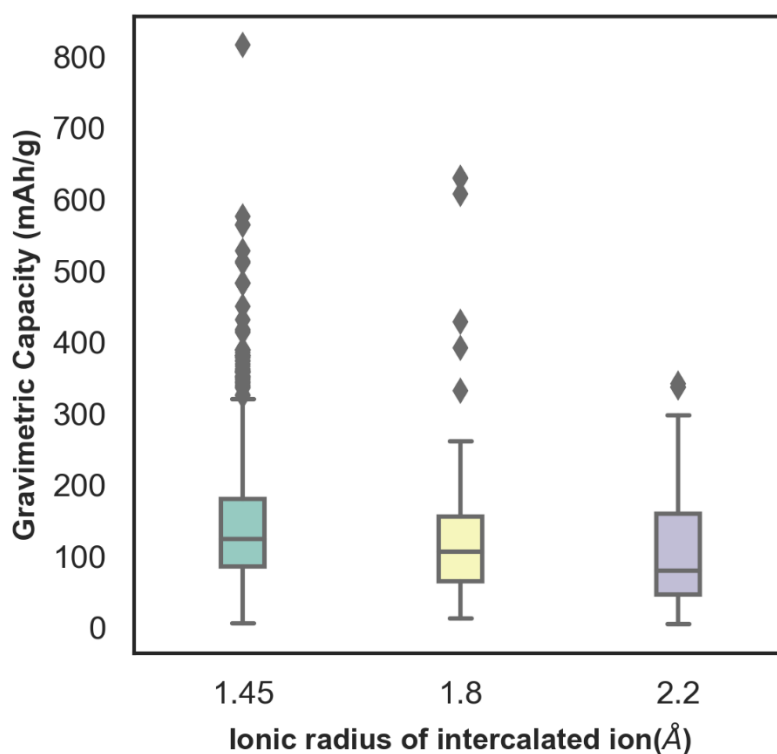


Figure 7: Range of capacity of Li, Na and K electrode materials.

Since the target variable capacity varies rapidly with the slide change in the electrode materials so in order to understand the distribution of the electrode materials across the capacity, we have plotted the range of % electrodes across per 100 mAh/g intervals of capacity. The distribution of capacity has shown in Figure 8.

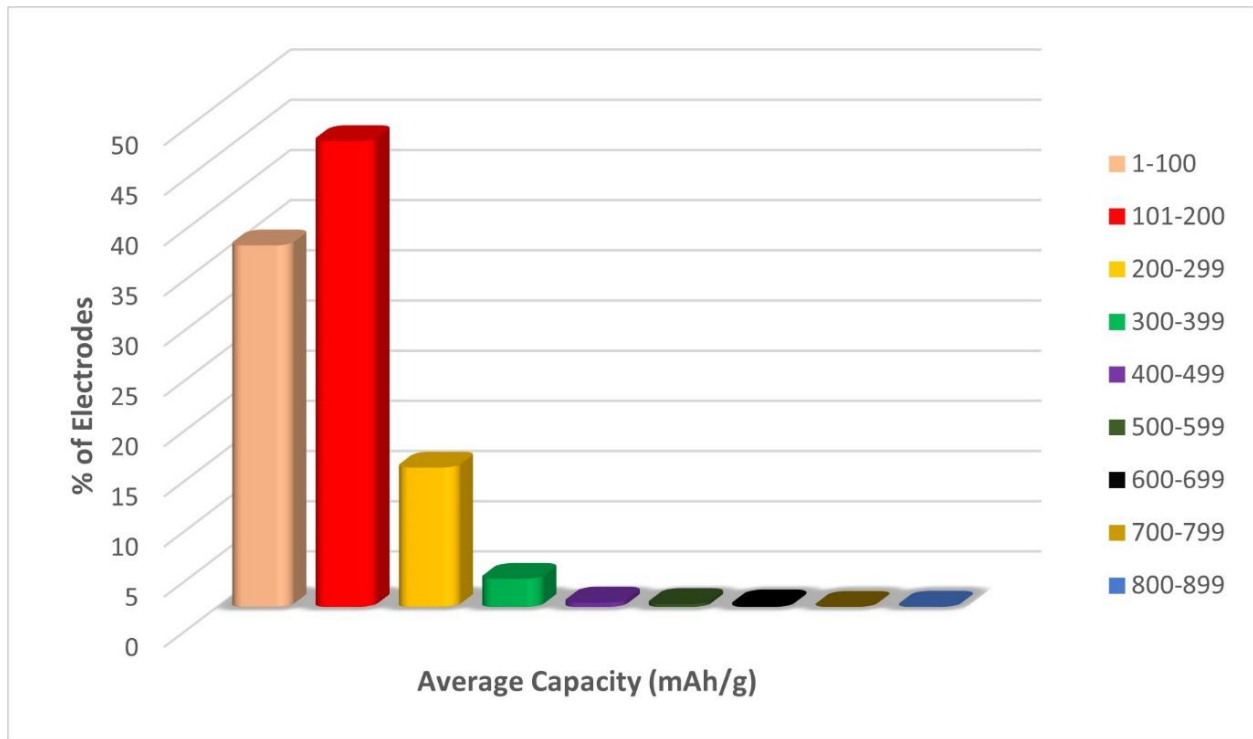


Figure 8: Distribution of capacity range across different electrode materials.

From the Figure 8, we can observe, more than 44% electrodes are there having capacity 101 to 200 mAh/g, around 35% electrodes having capacity 1 to 100 mAh/g and 15% electrodes having capacity 200 to 299 mAh/g. % of electrodes having capacity greater than 299 mAh/g is very less compared to the first three group of electrodes materials. Therefore, the sampling of target variable is not homogeneous rather highly heterogeneous, and this may be cause in a misinterpretation of the nature data by machine as there is a high chance of overestimate or remembering capacity data having range from 1 to 299. To avoid this overestimation, we have fit the ML models in three

different data set Na+K, Li, and Li+Na+K which has been discussed later. Further, the dimensionality reduction was done by the principal component analysis (PCA) on the selected descriptors to avoid the dimensionality curse. The PCA analysis shrink the dimension of the dataset by X %. We have assessed the performance of different ML models with PCA data and Without PCA data and it has been observed that the performance of ML models does not vary by large margin. However, since the dimension of the data getting reduced by the PCA method, we go ahead with the PCA data. The dimensionality reduction plot by PCA has been shown in Figure 9. From the plot it is very clear that the explained variance does not changing after the 37th number principal components which indicates that the first 37 principal components are enough to represent the whole data and we can remove the unnecessary principal components from the list.

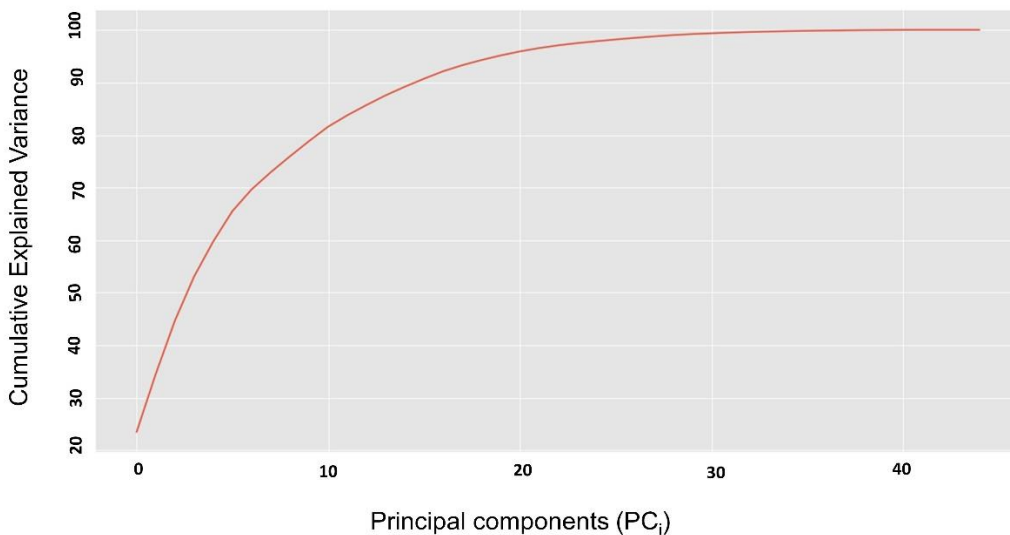


Figure 9: Change of explained variance with each principal components.

We have also calculated the contribution of each principal components (PC_i, $i = \{1, 45\}$) in terms of total variance and from the Figure 10 we have remove those principal components having negligible contribution.

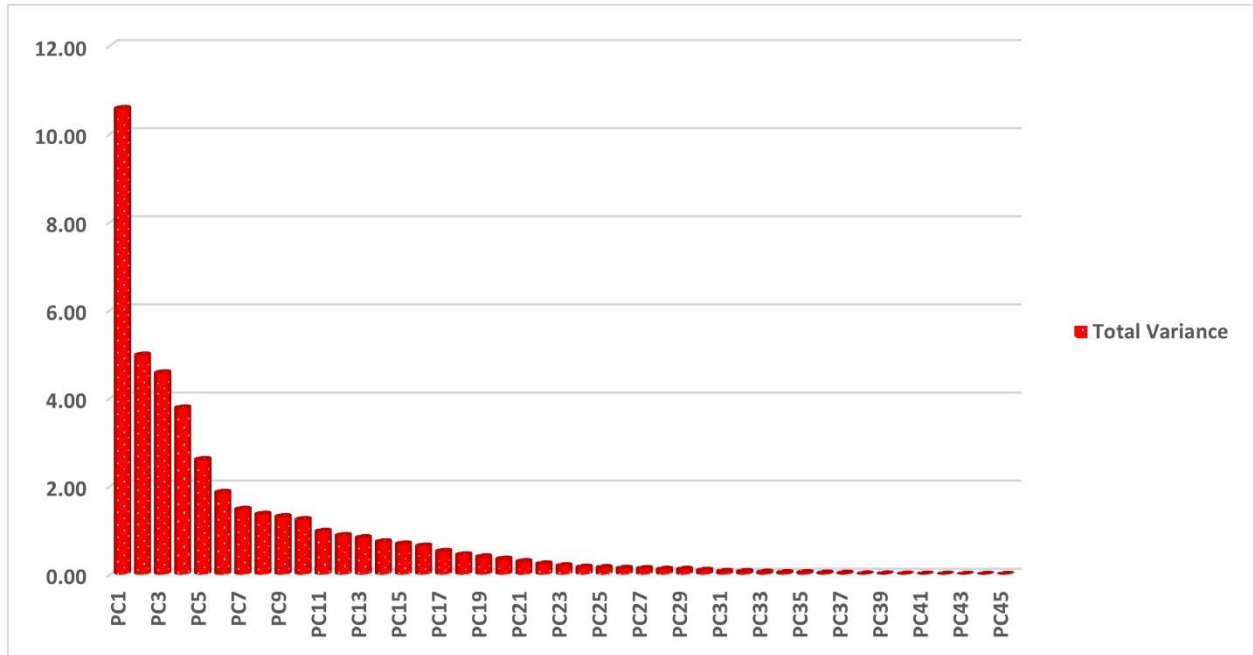


Figure 10: Contribution of each principal components in terms of total variance.

3. Results and Discussion

The analysis of our data set begins with fixed the target variable capacity. The overall data set splits in two set, train set composed of 80% of data and validation set composed of 20% data. Here we have compared three different machine learning algorithm Support vector machine (SVM), ExtraTress Regressor (EXR) and kernel ridge regression (KRR).

Since we are predicting continues value via ML, therefore it's belongs to a regression problem and to counter that we have used Support vector regression (SVR) a sub part of SVM. C the penalty term and gamma are two important parameters which needs to be optimize before fittings of SVR model. We have also tested our training data fixed in different kernel function of SVR like linear function, radial basis function (RBF) and polynomial function and then select the most optimized kernel considering the loss function as mean absolute percentage error (MAPE). As we have discussed before the large contribution of LIBs in overall dataset might result in a mimic of Li

data, we divided our data set in three sets Na+K, Li, and Li+Na+K dataset and then assessed the performance of ML models. The training set folded in 10-fold and for each cross-validation test 9-fold used for the training whether the remain 1-fold used for the assessment of the model performance in terms of MAPE as loss function. We have also calculated the standard deviation for each 10-fold cross validation set. By the cross validation test we have tried to be sampling our data in such a way so that machine can't overfit certain data which is the main reason behind a good train score but a very bad test score. The testing of different kernels are shown in Table 1. From the Table 1 it is observed that among different kernels RBF kernel function fitted well with less error as we assess our SVR model performance by checking the cross-validation score (cv_i).

Table 1: 10-fold cross validation (cv_i) test on full data set (Li+Na+K) having different kernel of Support vector regression (SVR).

SVR	cv_1	cv_2	cv_3	cv_4	cv_5	cv_6	cv_7	cv_8	cv_9	cv_{10}	SD	Mean MAPE	$MAPE_v$
Linear	0.62	0.36	0.37	0.49	0.46	0.38	0.32	0.29	0.29	0.36	0.102	0.395	0.314
RBF	0.46	0.36	0.32	0.33	0.35	0.40	0.23	0.19	0.20	0.35	0.08	0.32	0.236
Polynomial	0.60	0.39	0.64	0.55	0.46	0.44	0.44	0.31	0.43	0.53	0.099	0.479	0.396

The cross validation has been also performed for all three different dataset and shown in Table 1. The Mean MAPE shows the error on training set whereas $MAPE_v$ shows the error on validation set. For all the three dataset the similar trend has observed with respect to train error and validation error. Though it is expected lesser error for Na+K data as it has lowest number of data point however, the less error for Li+Na+K data set compared to other two dataset proves a better sampling of the data. The standard deviation for the dataset having higher number of data is low

compared to dataset having lower number of data which indicates that in overall full dataset the main deviation or outlier mainly arises from the Na+K data and not from the Li data.

Table 2: 10-fold cross validation (cv_i) test on three different dataset (Li+Na+K, Na+K, Li) having RBF kernel of Support vector regression (SVR).

SVR(RBF)	cv_1	cv_2	cv_3	cv_4	cv_5	cv_6	cv_7	cv_8	cv_9	cv_{10}	SD	Mean MAPE	$MAPE_v$
Na+K (C=75, gamma = 0.01)	0.54	0.50	0.25	0.38	0.24	0.18	0.12	0.19	0.18	0.23	0.14	0.282	0.284
Li (C=100, gamma=0.05)	0.44	0.36	0.35	0.31	0.40	0.32	0.25	0.21	0.19	0.34	0.08	0.320	0.257
Li+Na+K (C=100, gamma=0.05)	0.46	0.36	0.32	0.33	0.35	0.40	0.23	0.19	0.20	0.35	0.08	0.32	0.236

The hyperparameter tuning on C and gamma for SVM ML model has shown in Figure 11. It has been observed the best parameters for Li+Na+K and Li data are coming same whereas for Na+K data are different. After finding out the best hyperparameters for three different data set we have fitted SVM model for training on the training set and then validate the model in terms of MAPE utilizing the validation set.

The comparison between DFT calculated capacity and ML Predicted capacity has shown in the Figure 12. We have plotted the DFT calculated capacity vs ML predicted capacity using the best hyperparameters of SVM for all three different datasets.

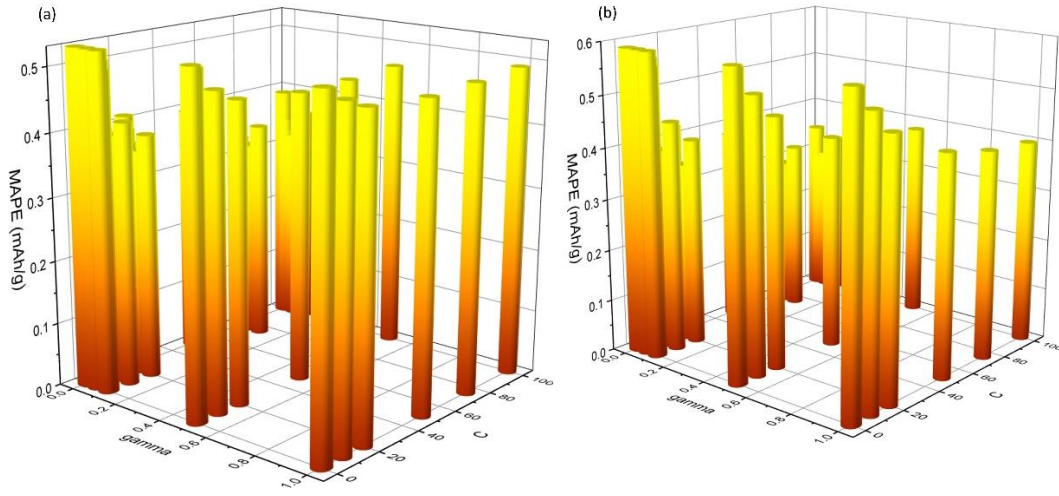


Figure 11: (a) Tuning of C and gamma parameter for Li+Na+K data set for SVM ML mode. (b) Tuning of C and gamma parameter for Na+K data for SVM ML model.

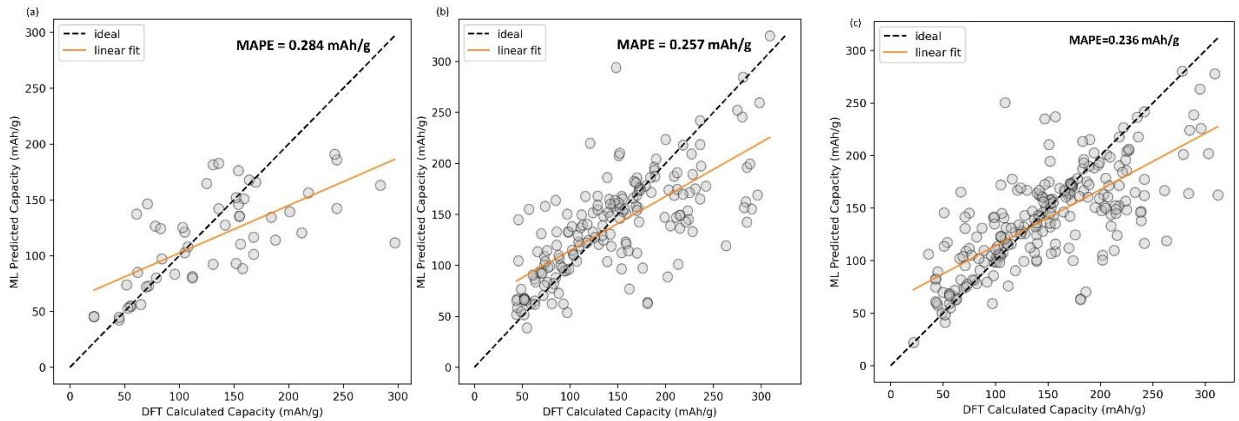


Figure 12: (a) Comparison between ML predicted capacity and DFT calculated capacity for SVM ML model using RBF kernel, $C=75$, $\gamma=0.01$ hyperparameters on Na+K dataset. (b) comparison between ML predicted capacity and DFT calculated capacity for SVM ML model using RBF kernel, $C=100$, $\gamma=0.05$ hyperparameters on Li dataset. (c) comparison between ML predicted capacity and DFT calculated capacity for SVM ML model using RBF kernel, $C = 100$, $\gamma= 0.05$ hyperparameters on Li+Na+K dataset.

Similarly, we have fitted our dataset in a tree-based ML model, ExtraTrees regressor (EXR). As we have compared the performance of SVM model in three different set here also we have plotted the same plot using EXR ML model. The number of trees and other parameters are optimized before fitting the EXR ML model. However, the optimized parameters remain same for all three different datasets. The error trend came same as the SVM model all the Li+Na+K data has given less error compared to others two data set. Moreover, the overall performance of EXR algorithm is better compared to the SVR algorithm.

Table 3: Cross validation score (cv_i), standard deviation (SD), Mean MAPE on training set and MAPE on validation set using EXR ML model.

EXR	cv_1	cv_2	cv_3	cv_4	cv_5	cv_6	cv_7	cv_8	cv_9	cv_{10}	SD	Mean MAPE	MAPE _v
Na+K	0.60	0.31	0.16	0.24	0.17	0.16	0.14	0.16	0.21	0.25	0.14	0.24	0.284
Li	0.43	0.37	0.32	0.36	0.37	0.31	0.21	0.20	0.23	0.38	0.08	0.32	0.257
Li+Na+K	0.47	0.33	0.26	0.32	0.34	0.39	0.23	0.19	0.25	0.33	0.08	0.31	0.236

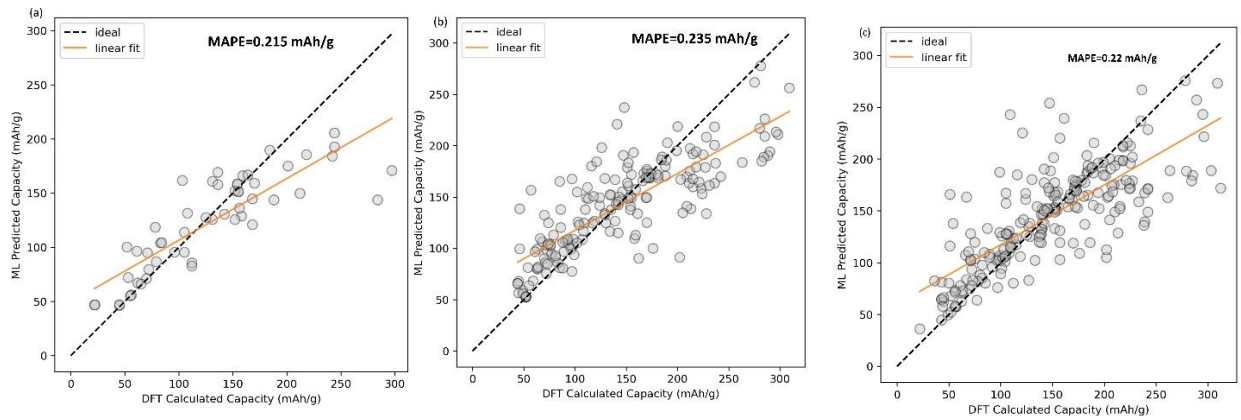


Figure 13: Comparison between ML predicted capacity and DFT calculated capacity for EXR ML

model number of trees=800, min_samples_leaf=3, min_samples_split=2 hyperparameters on (a) Na+K dataset. (b) Li dataset. (c) Li+Na+K dataset.

Furthermore, KRR has used for the fitting of the data where we have again checked 10-fold cross validation result after choosing the optimized hyperparameters. The result of 10-fold cross validation test has shown in Table 4. Among all these three ML algorithm KRR has fitted the Na+K data well compare to the others having $MAPE_v$ 0.153 mAh/g. Gamma and alpha are two important parameters for KRR algorithm. The optimization of these parameters is shown in Figure 14.

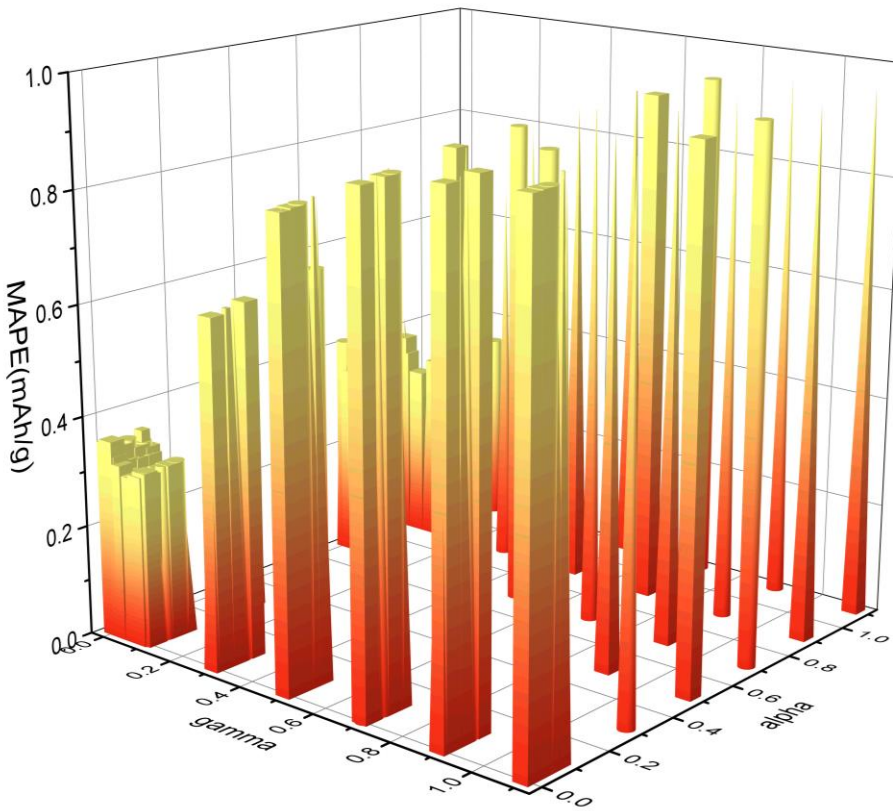


Figure 14: Optimization of gamma and alpha for KRR for Li+Na+K dataset.

Table 4: MAPE distribution of capacity, standard deviation (SD), Mean MAPE on training set and MAPE on validation set ($MAPE_v$) for 10 folds of training (cv_i) in KRR ML model trained with Na+K, Li, Li+Na+K data.

KRR	cv_1	cv_2	cv_3	cv_4	cv_5	cv_6	cv_7	cv_8	cv_9	cv_{10}	SD	Mean MAPE	$MAPE_v$
Na+K	0.50	0.19	0.16	0.19	0.12	0.14	0.19	0.19	0.13	0.24	0.108	0.207	0.153
Li	0.42	0.33	0.31	0.33	0.36	0.32	0.24	0.20	0.22	0.38	0.072	0.311	0.239
Li+Na+K	0.40	0.30	0.22	0.32	0.34	0.40	0.25	0.18	0.23	0.33	0.073	0.298	0.208

The DFT calculated capacity and ML predicted capacity for three different datasets has shown in Figure 15. Though the trend in training error and validation error in KRR is slightly different from the SVM and EXR ML model, overall, the performance of KRR ML model is better than the rest of two as KRR is able to mimic the nature of Na+K data better which is more important than to mimic Li ion data as our goal to predict the capacity for K ion battery. Therefore, from overall analysis on different dataset we can say KRR perform better.

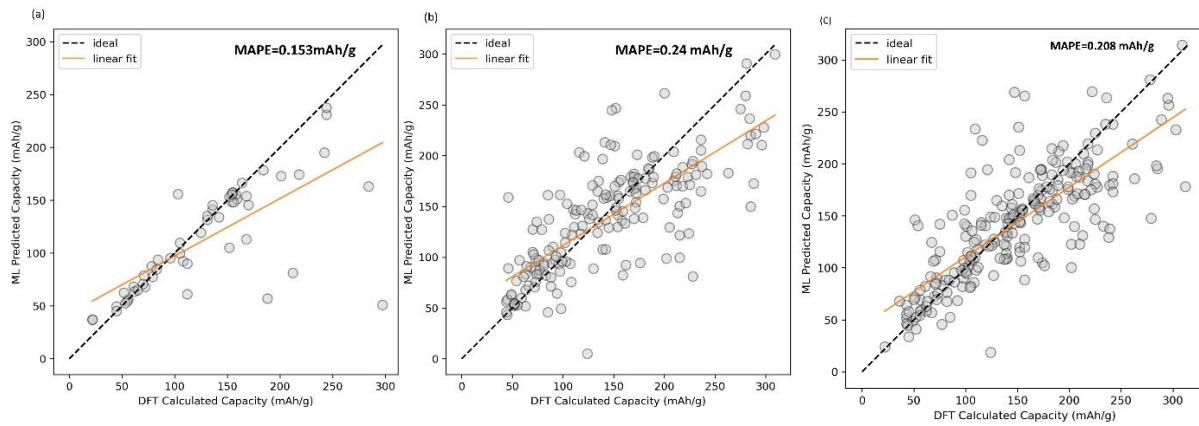


Figure 15: Comparison between ML predicted capacity and DFT calculated capacity for KRR ML model (kernel=Laplacian, alpha=0.024239, gamma=0.047051, degree=2 hyperparameters) on (a) Na+K dataset. (b) Li dataset. (c) Li+Na+K dataset. We have also fitted random forest regressor (RFR). The best hyperparameters of RFR and optimized number of trees is attached with Text S1 and Figure S1 respectively and the cross validation score is attached with Table S2. Optimized hyperparameters and mean absolute percentage error (MAPE) for decision trees regressor has also given in Table S3.

4. DFT Validation

To validate our calculated capacity for various electrode materials, we have considered five different electrode materials and checked their maximum specific capacity by carrying out first principles calculations using the projector augmented wave (PAW) method as implemented in the Vienna Ab initio Simulation Package.^{63–68} Moreover, the generalized gradient approximation of Perdew–Burke–Ernzerhof (GGA-PBE) has been considered as the exchange correlation potentials and the energy cutoff is set to 470 eV. Furthermore, the dispersion energy corrections have been considered by incorporating DFT-D3 method of Grimme.^{69,70} All of the structures are relaxed until the Hellmann–Feynman force criteria of $< 0.01 \text{ eV/\AA}$ and the total energy convergence criteria of 10^{-4} eV .

The five unintercalated materials we have considered are Mn_4NiO_8 , FeO_2 , $\text{Fe}(\text{CoO}_3)_2$, V_5O_{12} and CoPO_4 . The structures of the system have been taken from Materials Project database. The fully intercalated systems with maximum capacity are represented in Figure 14.

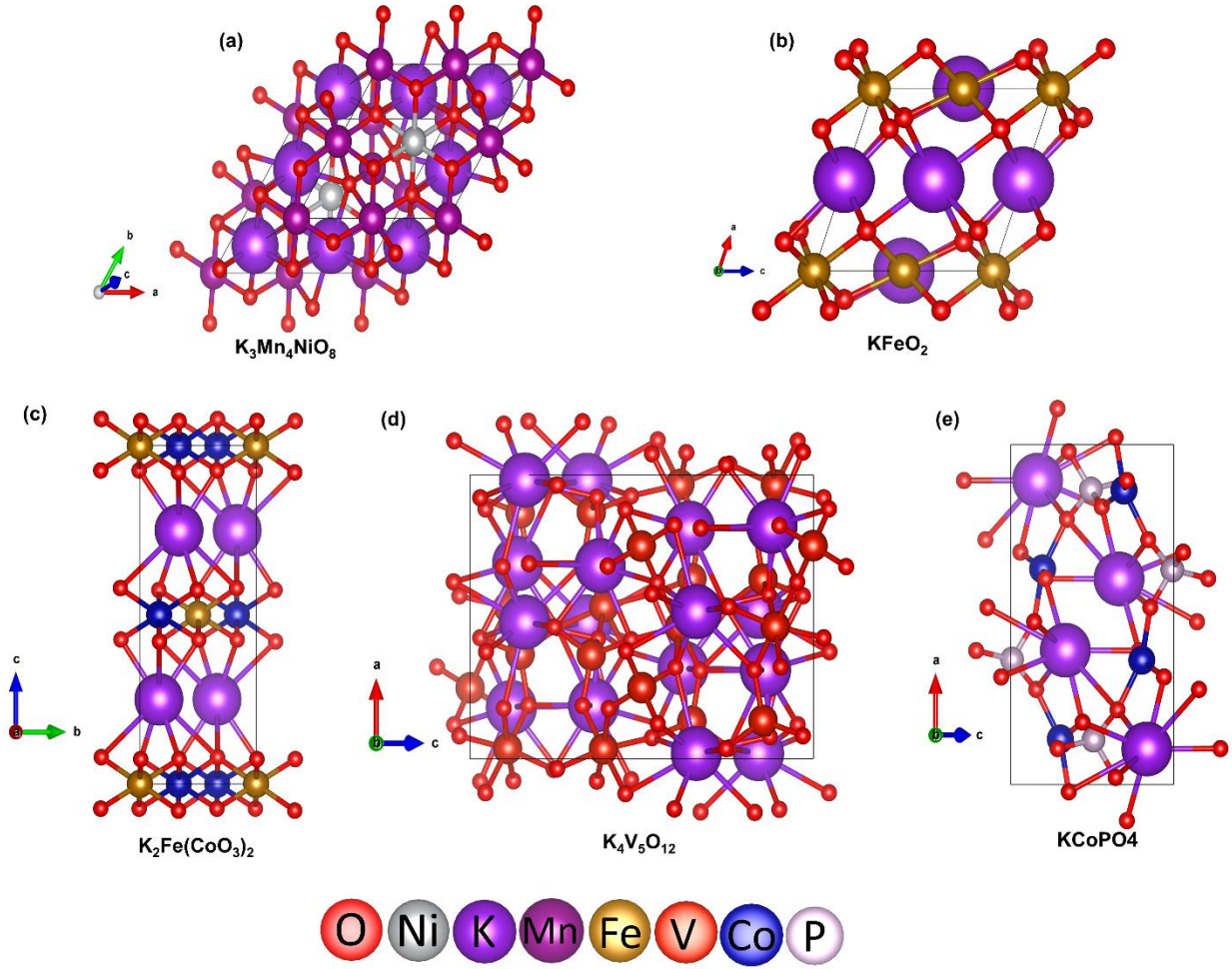


Figure 16: DFT optimized structures of K intercalated electrode materials.

Using the value of specific capacity from ML results, we obtained the number of intercalating ions using the equation,

$$C = \frac{zxF}{M_f}$$

where z represents the charge on intercalating ions (1 in case of K), x represents the number of intercalating ions and F is the Faraday constant (26.8 Ah mol^{-1}). M_f represents the molecular weight of the formula unit of the electrode material. The comparison between the ML predicted

data and that of DFT are presented in Table 7. Thus we can see that our ML results are very close to the independent DFT results and can be used as alternative for time consuming DFT calculations in general.

Table 5: Comparision between number of K ion intercalating calcaulation by machine learning and DFT.

Electrode materials	Predicted data	DFT calculated data
	No. of intercalating K ions	No. of intercalating K ions
Mn_4NiO_8	3.5	3
FeO_2	0.7	1
$\text{Fe}(\text{CoO}_3)_2$	2.2	2
V_5O_{12}	3.6	4
CoPO_4	1.2	1

5. Conclusion

In this work we have predicted specific capacity of prospective K-ion battery electrode materials. We have considered Li, Na and K-ion electrode materials and their available battery data from Materials Project database. Suitable features have been considered and developed to train the various machine learning algorithms. The available data has been divided into train set and validation set. The train set has been fitted using various ML algorithms like support vector machine, extratrees regressor and kernel ridge regression to learn the nature of the data and features. Some statistical methods of data analysis like box plot for outlier detection, joint plot to understand the distribution of descriptors, heatmap for the correlation metrics, principal component analysis to reduce the dimension of our dataset have been utilized. We have evaluated the performance of considered machine learning models by compairing the mean absolute percentage error between

training set and validation set in each case. Further, adopting kernel ridge regression we have predicted the capacity of unknown electrode materials for K-ion battery (Table S4). Using the value of specific capacity, the number of intercalated K ions in the formula unit of the non-intercalated electrode material compounds have been calculated. DFT calculations have been performed for sample electrode materials to verify that our ML model can give similar results. Due to large range in output specific capacity values, the ML techniques are still not able to provide very accurate results and further improvements are expected in future. Thus, implementing ML approach is much more faster compared to the computationally demanding quantum mechanical methods for quick screening of electrode materials which will help to guide the experiments for developing electrode materials for metal ion batteries.

6. Supporting Information

The supporting information contents are elemental properties to generate choice-based feature vectorization, heatmap showing the intercorrelation among different selected features, best hyperparameters found for random forest regressor (RFR), estimation of optimized number of trees for random forest ML model, cross validation score for random forest regressor, optimized hyperparameters and mean absolute percentage error for decision trees regressor, predicted capacity and number of K ion intercalated per unit formula weight.

7. Acknowledgements

We thank IIT Indore for lab and computing facilities. This work is supported by DST-SERB (Project Number CRG/2018/001131), SPARC (Project Number SPARC/2018-2019/P116/SL) New Delhi, and CSIR (project 01(3046)/21/EMR-II). S.M. and S.D. thank CSIR for research fellowship. D.R. thanks MHRD for research fellowship.

8. References

- (1) Yang, Z.; Zhang, J.; Kintner-Meyer, M. C. W.; Lu, X.; Choi, D.; Lemmon, J. P.; Liu, J. Electrochemical Energy Storage for Green Grid. *Chemical Reviews* **2011**, *111* (5), 3577–3613. <https://doi.org/10.1021/CR100290V>.
- (2) Nitta, N.; Wu, F.; Lee, J. T.; Yushin, G. Li-Ion Battery Materials: Present and Future. *Materials Today* **2015**, *18* (5), 252–264. <https://doi.org/10.1016/J.MATTOD.2014.10.040>.
- (3) Winter, M.; Barnett, B.; Xu, K. Before Li Ion Batteries. *Chemical Reviews* **2018**, *118* (23), 11433–11456. <https://doi.org/10.1021/ACS.CHEMREV.8B00422>.
- (4) Dunn, B.; Kamath, H.; Tarascon, J. M. Electrical Energy Storage for the Grid: A Battery of Choices. *Science* **2011**, *334* (6058), 928–935. https://doi.org/10.1126/SCIENCE.1212741/SUPPL_FILE/DUNN-SOM.PDF.
- (5) Cheng, F.; Liang, J.; Tao, Z.; Chen, J.; Cheng, F. Y.; Liang, J.; Tao, Z. L.; Chen, J. Functional Materials for Rechargeable Batteries. *Advanced Materials* **2011**, *23* (15), 1695–1715. <https://doi.org/10.1002/ADMA.201003587>.
- (6) Tarascon, J. M. Is Lithium the New Gold? *Nature Chemistry* **2010**, *2* (6), 510–510. <https://doi.org/10.1038/nchem.680>.
- (7) Joshi, R. P.; Ozdemir, B.; Barone, V.; Peralta, J. E. Hexagonal BC₃: A Robust Electrode Material for Li, Na, and K Ion Batteries. *Journal of Physical Chemistry Letters* **2015**, *6* (14), 2728–2732. https://doi.org/10.1021/ACS.JPCLETT.5B01110/SUPPL_FILE/JZ5B01110_SI_001.PDF.
- (8) Bhauriyal, P.; Mahata, A.; Pathak, B. Hexagonal BC₃ Electrode for a High-Voltage Al-Ion Battery. *Journal of Physical Chemistry C* **2017**, *121* (18), 9748–9756. https://doi.org/10.1021/ACS.JPCC.7B02290/SUPPL_FILE/JP7B02290_SI_001.PDF.
- (9) Posada, J. O. G.; Rennie, A. J. R.; Villar, S. P.; Martins, V. L.; Marinaccio, J.; Barnes, A.; Glover, C. F.; Worsley, D. A.; Hall, P. J. Aqueous Batteries as Grid Scale Energy Storage Solutions. *Renewable and Sustainable Energy Reviews* **2017**, *68*, 1174–1182. <https://doi.org/10.1016/J.RSER.2016.02.024>.
- (10) Liu, K.; Liu, Y.; Lin, D.; Pei, A.; Cui, Y. Materials for Lithium-Ion Battery Safety. *Science Advances* **2018**, *4* (6). <https://doi.org/10.1126/SCIADV.AAS9820/ASSET/8682E3B0-F43B-4EBF-A50B-AE67A9ACFF5C/ASSETS/GRAPHIC/AAS9820-F6.JPEG>.
- (11) Tarascon, J. M.; Armand, M. Issues and Challenges Facing Rechargeable Lithium Batteries. *Materials for Sustainable Energy: A Collection of Peer-Reviewed Research and Review Articles from Nature Publishing Group* **2010**, 171–179. https://doi.org/10.1142/9789814317665_0024.
- (12) Nithya, C.; Gopukumar, S. Sodium Ion Batteries: A Newer Electrochemical Storage. *Wiley Interdisciplinary Reviews: Energy and Environment* **2015**, *4* (3), 253–278. <https://doi.org/10.1002/WENE.136>.
- (13) Larcher, D.; Tarascon, J. M. Towards Greener and More Sustainable Batteries for Electrical Energy Storage. *Nature Chemistry* **2014**, *7* (1), 19–29. <https://doi.org/10.1038/nchem.2085>.

- (14) Scrosati, B.; Garche, J. Lithium Batteries: Status, Prospects and Future. *Journal of Power Sources* **2010**, *195* (9), 2419–2430. <https://doi.org/10.1016/J.JPOWSOUR.2009.11.048>.
- (15) Zhang, W.; Liu, Y.; Guo, Z. Approaching High-Performance Potassium-Ion Batteries via Advanced Design Strategies and Engineering. *Science Advances* **2019**, *5* (5). <https://doi.org/10.1126/SCIADV.AAV7412/ASSET/5FC5F123-66F7-4FE1-9292-4F533B36083F/ASSETS/GRAPHIC/AAV7412-F6.JPEG>.
- (16) Kirkpatrick, P.; Ellis, C. Chemical Space. *Nature* **2004**, *432* (7019), 823. <https://doi.org/10.1038/432823A>.
- (17) von Lilienfeld, O. A. Quantum Machine Learning in Chemical Compound Space. *Angewandte Chemie International Edition* **2018**, *57* (16), 4164–4169. <https://doi.org/10.1002/ANIE.201709686>.
- (18) Mullard, A. The Drug-Maker's Guide to the Galaxy. *Nature* **2017**, *549* (7673), 445–447. <https://doi.org/10.1038/549445A>.
- (19) Draxl, C.; Scheffler, M. NOMAD: The FAIR Concept for Big Data-Driven Materials Science. *MRS Bulletin 2018 43:9* **2018**, *43* (9), 676–682. <https://doi.org/10.1557/MRS.2018.208>.
- (20) Saal, J. E.; Kirklin, S.; Aykol, M.; Meredig, B.; Wolverton, C. Materials Design and Discovery with High-Throughput Density Functional Theory: The Open Quantum Materials Database (OQMD). *JOM* **2013**, *65* (11), 1501–1509. <https://doi.org/10.1007/S11837-013-0755-4/FIGURES/7>.
- (21) Kirklin, S.; Saal, J. E.; Meredig, B.; Thompson, A.; Doak, J. W.; Aykol, M.; Rühl, S.; Wolverton, C. The Open Quantum Materials Database (OQMD): Assessing the Accuracy of DFT Formation Energies. *npj Computational Materials 2015 1:1* **2015**, *1* (1), 1–15. <https://doi.org/10.1038/npjcompumats.2015.10>.
- (22) Curtarolo, S.; Setyawan, W.; Hart, G. L. W.; Jahnatek, M.; Chepulskii, R. v.; Taylor, R. H.; Wang, S.; Xue, J.; Yang, K.; Levy, O.; Mehl, M. J.; Stokes, H. T.; Demchenko, D. O.; Morgan, D. AFLOW: An Automatic Framework for High-Throughput Materials Discovery. *Computational Materials Science* **2012**, *58*, 218–226. <https://doi.org/10.1016/J.COMMATSCI.2012.02.005>.
- (23) Gossett, E.; Toher, C.; Oses, C.; Isayev, O.; Legrain, F.; Rose, F.; Zurek, E.; Carrete, J.; Mingo, N.; Tropsha, A.; Curtarolo, S. AFLOW-ML: A RESTful API for Machine-Learning Predictions of Materials Properties. *Computational Materials Science* **2017**, *152*, 134–145. <https://doi.org/10.1016/j.commatsci.2018.03.075>.
- (24) 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. Commentary: The Materials Project: A Materials Genome Approach to Accelerating Materials Innovation. *APL Materials* **2013**, *1* (1), 011002. <https://doi.org/10.1063/1.4812323>.
- (25) Ong, S. P.; Cholia, S.; Jain, A.; Brafman, M.; Gunter, D.; Ceder, G.; Persson, K. A. The Materials Application Programming Interface (API): A Simple, Flexible and Efficient API for Materials Data Based on REpresentational State Transfer (REST) Principles. *Computational Materials Science* **2015**, *97*, 209–215. <https://doi.org/10.1016/J.COMMATSCI.2014.10.037>.

- (26) Joshi, R. P.; Trepte, K.; Withanage, K. P. K.; Sharkas, K.; Yamamoto, Y.; Basurto, L.; Zope, R. R.; Baruah, T.; Jackson, K. A.; Peralta, J. E. Fermi-Löwdin Orbital Self-Interaction Correction to Magnetic Exchange Couplings. *The Journal of Chemical Physics* **2018**, *149* (16), 164101. <https://doi.org/10.1063/1.5050809>.
- (27) Kaloni, T. P.; Joshi, R. P.; Adhikari, N. P.; Schwingenschlögl, U. Band Gap Tuning in BN-Doped Graphene Systems with High Carrier Mobility. *Applied Physics Letters* **2014**, *104* (7), 073116. <https://doi.org/10.1063/1.4866383>.
- (28) Ramprasad, R.; Batra, R.; Pilania, G.; Mannodi-Kanakkithodi, A.; Kim, C. Machine Learning in Materials Informatics: Recent Applications and Prospects. *npj Computational Materials* **2017** *3*:1 **2017**, *3* (1), 1–13. <https://doi.org/10.1038/s41524-017-0056-5>.
- (29) Bassman, L.; Rajak, P.; Kalia, R. K.; Nakano, A.; Sha, F.; Sun, J.; Singh, D. J.; Aykol, M.; Huck, P.; Persson, K.; Vashishta, P. Active Learning for Accelerated Design of Layered Materials. *npj Computational Materials* **2018** *4*:1 **2018**, *4* (1), 1–9. <https://doi.org/10.1038/s41524-018-0129-0>.
- (30) Zhang, Y.; Ling, C. A Strategy to Apply Machine Learning to Small Datasets in Materials Science. *npj Computational Materials* **2018** *4*:1 **2018**, *4* (1), 1–8. <https://doi.org/10.1038/s41524-018-0081-z>.
- (31) Butler, K. T.; Davies, D. W.; Cartwright, H.; Isayev, O.; Walsh, A. Machine Learning for Molecular and Materials Science. *Nature* **2018** *559*:7715 **2018**, *559* (7715), 547–555. <https://doi.org/10.1038/s41586-018-0337-2>.
- (32) Schleder, G. R.; Padilha, A. C. M.; Acosta, C. M.; Costa, M.; Fazzio, A. From DFT to Machine Learning: Recent Approaches to Materials Science—a Review. *Journal of Physics: Materials* **2019**, *2* (3), 032001. <https://doi.org/10.1088/2515-7639/AB084B>.
- (33) Dong, Y.; Wu, C.; Zhang, C.; Liu, Y.; Cheng, J.; Lin, J. Bandgap Prediction by Deep Learning in Configurationally Hybridized Graphene and Boron Nitride. *npj Computational Materials* **2019** *5*:1 **2019**, *5* (1), 1–8. <https://doi.org/10.1038/s41524-019-0165-4>.
- (34) Zhuo, Y.; Mansouri Tehrani, A.; Brgoch, J. Predicting the Band Gaps of Inorganic Solids by Machine Learning. *Journal of Physical Chemistry Letters* **2018**, *9* (7), 1668–1673. https://doi.org/10.1021/ACS.JPCLETT.8B00124/SUPPL_FILE/JZ8B00124_SI_003.XLSX.
- (35) Kolb, B.; Lentz, L. C.; Kolpak, A. M. Discovering Charge Density Functionals and Structure-Property Relationships with PROPhet: A General Framework for Coupling Machine Learning and First-Principles Methods. *Scientific Reports* **2017** *7*:1 **2017**, *7* (1), 1–9. <https://doi.org/10.1038/s41598-017-01251-z>.
- (36) Pilania, G.; Wang, C.; Jiang, X.; Rajasekaran, S.; Ramprasad, R. Accelerating Materials Property Predictions Using Machine Learning. *Scientific Reports* **2013** *3*:1 **2013**, *3* (1), 1–6. <https://doi.org/10.1038/srep02810>.
- (37) Pilania, G.; Mannodi-Kanakkithodi, A.; Uberuaga, B. P.; Ramprasad, R.; Gubernatis, J. E.; Lookman, T. Machine Learning Bandgaps of Double Perovskites. *Scientific Reports* **2016** *6*:1 **2016**, *6* (1), 1–10. <https://doi.org/10.1038/srep19375>.
- (38) Takahashi, K.; Takahashi, L.; Miyazato, I.; Tanaka, Y. Searching for Hidden Perovskite Materials for Photovoltaic Systems by Combining Data Science and First Principle Calculations. *ACS*

- Photonics* **2018**, *5* (3), 771–775.
https://doi.org/10.1021/ACSPHOTONICS.7B01479/SUPPL_FILE/PH7B01479_SI_001.PDF.
- (39) Sodeyama, K.; Igarashi, Y.; Nakayama, T.; Tateyama, Y.; Okada, M. Liquid Electrolyte Informatics Using an Exhaustive Search with Linear Regression. *Physical Chemistry Chemical Physics* **2018**, *20* (35), 22585–22591. <https://doi.org/10.1039/C7CP08280K>.
 - (40) Okamoto, Y.; Kubo, Y. Ab Initio Calculations of the Redox Potentials of Additives for Lithium-Ion Batteries and Their Prediction through Machine Learning. *ACS Omega* **2018**, *3* (7), 7868–7874. https://doi.org/10.1021/ACSOMEGA.8B00576/SUPPL_FILE/AO8B00576_SI_001.XLSX.
 - (41) Jalem, R.; Nakayama, M.; Kasuga, T. An Efficient Rule-Based Screening Approach for Discovering Fast Lithium Ion Conductors Using Density Functional Theory and Artificial Neural Networks. *Journal of Materials Chemistry A* **2013**, *2* (3), 720–734.
<https://doi.org/10.1039/C3TA13235H>.
 - (42) Fujimura, K.; Seko, A.; Koyama, Y.; Kuwabara, A.; Kishida, I.; Shitara, K.; Fisher, C. A. J.; Moriwake, H.; Tanaka, I. Accelerated Materials Design of Lithium Superionic Conductors Based on First-Principles Calculations and Machine Learning Algorithms. *Advanced Energy Materials* **2013**, *3* (8), 980–985. <https://doi.org/10.1002/AENM.201300060>.
 - (43) Kireeva, N.; Pervov, V. S. Materials Space of Solid-State Electrolytes: Unraveling Chemical Composition–Structure–Ionic Conductivity Relationships in Garnet-Type Metal Oxides Using Cheminformatics Virtual Screening Approaches. *Physical Chemistry Chemical Physics* **2017**, *19* (31), 20904–20918. <https://doi.org/10.1039/C7CP00518K>.
 - (44) 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. *The Journal of Chemical Physics* **2019**, *150* (21), 214701. <https://doi.org/10.1063/1.5093220>.
 - (45) Sendek, A. D.; Yang, Q.; Cubuk, E. D.; Duerloo, K. A. N.; Cui, Y.; Reed, E. J. Holistic Computational Structure Screening of More than 12 000 Candidates for Solid Lithium-Ion Conductor Materials. *Energy & Environmental Science* **2017**, *10* (1), 306–320.
<https://doi.org/10.1039/C6EE02697D>.
 - (46) Schütt, K. T.; Glawe, H.; Brockherde, F.; Sanna, A.; Müller, K. R.; Gross, E. K. U. How to Represent Crystal Structures for Machine Learning: Towards Fast Prediction of Electronic Properties. *Physical Review B - Condensed Matter and Materials Physics* **2014**, *89* (20), 205118.
<https://doi.org/10.1103/PHYSREVB.89.205118/FIGURES/3/MEDIUM>.
 - (47) Kitchin, J. R. Machine Learning in Catalysis. *Nature Catalysis* **2018**, *1* (4), 230–232.
<https://doi.org/10.1038/s41929-018-0056-y>.
 - (48) Goldsmith, B. R.; Esterhuizen, J.; Liu, J. X.; Bartel, C. J.; Sutton, C. Machine Learning for Heterogeneous Catalyst Design and Discovery. *AIChE Journal* **2018**, *64* (7), 2311–2323.
<https://doi.org/10.1002/AIC.16198>.
 - (49) Seko, A.; Hayashi, H.; Nakayama, K.; Takahashi, A.; Tanaka, I. Representation of Compounds for Machine-Learning Prediction of Physical Properties. *Physical Review B* **2017**, *95* (14), 144110.
<https://doi.org/10.1103/PHYSREVB.95.144110/FIGURES/5/MEDIUM>.

- (50) Seko, A.; Maekawa, T.; Tsuda, K.; Tanaka, I. Machine Learning with Systematic Density-Functional Theory Calculations: Application to Melting Temperatures of Single- and Binary-Component Solids. *Physical Review B - Condensed Matter and Materials Physics* **2014**, *89* (5), 054303. <https://doi.org/10.1103/PHYSREVB.89.054303/FIGURES/7/MEDIUM>.
- (51) Ward, L.; Agrawal, A.; Choudhary, A.; Wolverton, C. A General-Purpose Machine Learning Framework for Predicting Properties of Inorganic Materials. *npj Computational Materials* **2016**, *2*:1 **2016**, *2* (1), 1–7. <https://doi.org/10.1038/npjcompumats.2016.28>.
- (52) Faber, F. A.; Lindmaa, A.; von Lilienfeld, O. A.; Armiento, R. Machine Learning Energies of 2 Million Elpasolite (ABC2D6) Crystals. *Physical Review Letters* **2016**, *117* (13), 135502. <https://doi.org/10.1103/PHYSREVLETT.117.135502/FIGURES/2/MEDIUM>.
- (53) Deml, A. M.; O’Hayre, R.; Wolverton, C.; Stevanović, V. Predicting Density Functional Theory Total Energies and Enthalpies of Formation of Metal-Nonmetal Compounds by Linear Regression. *Physical Review B* **2016**, *93* (8), 085142. <https://doi.org/10.1103/PHYSREVB.93.085142/FIGURES/6/MEDIUM>.
- (54) Meredig, B.; Agrawal, A.; Kirklin, S.; Saal, J. E.; Doak, J. W.; Thompson, A.; Zhang, K.; Choudhary, A.; Wolverton, C. Combinatorial Screening for New Materials in Unconstrained Composition Space with Machine Learning. *Physical Review B - Condensed Matter and Materials Physics* **2014**, *89* (9), 094104. <https://doi.org/10.1103/PHYSREVB.89.094104/FIGURES/4/MEDIUM>.
- (55) Brockherde, F.; Vogt, L.; Li, L.; Tuckerman, M. E.; Burke, K.; Müller, K. R. Bypassing the Kohn-Sham Equations with Machine Learning. *Nature Communications* **2017**, *8*:1 **2017**, *8* (1), 1–10. <https://doi.org/10.1038/s41467-017-00839-3>.
- (56) Mills, K.; Spanner, M.; Tamblyn, I. Deep Learning and the Schrödinger Equation. *Physical Review A* **2017**, *96* (4), 042113. <https://doi.org/10.1103/PHYSREVA.96.042113/FIGURES/11/MEDIUM>.
- (57) Li, L.; Snyder, J. C.; Pelaschier, I. M.; Huang, J.; Niranjana, U. N.; Duncan, P.; Rupp, M.; Müller, K. R.; Burke, K. Understanding Machine-Learned Density Functionals. *International Journal of Quantum Chemistry* **2016**, *116* (11), 819–833. <https://doi.org/10.1002/QUA.25040>.
- (58) Wu, J.; Zhang, C.; Chen, Z. An Online Method for Lithium-Ion Battery Remaining Useful Life Estimation Using Importance Sampling and Neural Networks. *Applied Energy* **2016**, *173*, 134–140. <https://doi.org/10.1016/J.APENERGY.2016.04.057>.
- (59) Cheng, D.; Sha, W.; Wang, L.; Tang, S.; Ma, A.; Chen, Y.; Wang, H.; Lou, P.; Lu, S.; Cao, Y. C. Solid-State Lithium Battery Cycle Life Prediction Using Machine Learning. *Applied Sciences* **2021**, *Vol. 11, Page 4671* **2021**, *11* (10), 4671. <https://doi.org/10.3390/AP11104671>.
- (60) 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. *Computational Materials Science* **2013**, *68*, 314–319. <https://doi.org/10.1016/J.COMMATSCI.2012.10.028>.
- (61) Zhou, F.; Cococcioni, M.; Marianetti, C. A.; Morgan, D.; Ceder, G. First-Principles Prediction of Redox Potentials in Transition-Metal Compounds with LDA + U. *Physical Review B - Condensed*

Matter and Materials Physics **2004**, 70 (23), 1–8.

<https://doi.org/10.1103/PHYSREVB.70.235121/FIGURES/6/MEDIUM>.

- (62) Wang, A. Y. T.; Murdock, R. J.; Kauwe, S. K.; Oliynyk, A. O.; Gurlo, A.; Brgoch, J.; Persson, K. A.; Sparks, T. D. Machine Learning for Materials Scientists: An Introductory Guide toward Best Practices. *Chemistry of Materials* **2020**, 32 (12), 4954–4965.
https://doi.org/10.1021/ACS.CHEMMATER.0C01907/SUPPL_FILE/CM0C01907_SI_001.PDF.
- (63) Blöchl, P. E. Projector Augmented-Wave Method. *Physical Review B* **1994**, 50 (24), 17953.
<https://doi.org/10.1103/PhysRevB.50.17953>.
- (64) Kresse, G.; Joubert, D. From ultrasoft pseudopotentials to the projector augmented-wave method. *Physical Review B* **1999**, 59 (3), 1758. <https://doi.org/10.1103/PhysRevB.59.1758>.
- (65) Kresse, G.; Hafner, J. *Ab Initio* Molecular-Dynamics Simulation of the Liquid-Metal–Amorphous-Semiconductor Transition in Germanium. *Physical Review B* **1994**, 49 (20), 14251.
<https://doi.org/10.1103/PhysRevB.49.14251>.
- (66) Kresse, G.; Hafner, J. *Ab Initio* Molecular Dynamics for Liquid Metals. *Physical Review B* **1993**, 47 (1), 558. <https://doi.org/10.1103/PhysRevB.47.558>.
- (67) Kresse, G.; Furthmüller, J. Efficiency of *Ab-Initio* Total Energy Calculations for Metals and Semiconductors Using a Plane-Wave Basis Set. *Computational Materials Science* **1996**, 6 (1), 15–50. [https://doi.org/10.1016/0927-0256\(96\)00008-0](https://doi.org/10.1016/0927-0256(96)00008-0).
- (68) Kresse, G.; Furthmüller, J. Efficient iterative schemes for *Ab Initio* Total-Energy Calculations Using a Plane-Wave Basis Set. *Physical Review B* **1996**, 54 (16), 11169.
<https://doi.org/10.1103/PhysRevB.54.11169>.
- (69) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Physical Review Letters* **1996**, 77 (18), 3865. <https://doi.org/10.1103/PhysRevLett.77.3865>.
- (70) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A consistent and accurate *Ab Initio* parametrization of Density Functional Dispersion Correction (DFT-D) for the 94 elements H–Pu. *The Journal of Chemical Physics* **2010**, 132 (15), 154104. <https://doi.org/10.1063/1.3382344>.

TOC Graphics

