Machine Learning Model for Predicting Formation Energies for Lithium-ion Battery Materials

KM Monareng[1], RR Maphanga[2] and SP Ntoahae[1]

[1]Department of Physics, University of Limpopo, Private bag x 1106, Sovenga, 0727
[2]Next Generation Enterprises and Institutions, Council for Scientific and Industrial Research, P.O. Box 395, Pretoria, 0001

SAIP 22-29 JULY 2021

ABSTRACT

Machine learning methods have recently found applications in many areas of physics, chemistry, biology and materials science, where large datasets are available. In this paper, machine learning methods are used to predict the formation energies of lithium-ion battery (LIB) materials. Thus, using LIB materials' properties calculated from density functional theory as an input dataset, as well as feature vectors from properties of chemical compounds and elemental properties of their constituents, different machine learning algorithms are explored in order to predict the formation energies for the battery materials. Models based on different algorithms, i.e., extremely randomized trees, gradient boosting, light gradient boosting machine, catboost and random forest were developed and evaluated. The catboost regressor model was found to be the best model in predicting the formation energies, with accuracy of 0.95 and 0.06 for coefficient of determination and mean square error, respectively. Thus, the features used to predict the formation energies have predictive capability with a high accuracy.

INTRODUCTION

- LIBs have revolutionized the energy storage technology and played a vital role in transforming portable electronic devices, particularly in terms of lifetime, weight, size and performance [1].
- Despite being dominating energy source for portable devices, the transferability of conventional LIBs to higher energy scales remains a challenge owing to their relatively low energy density [2-4].
- In the recent past, computational modelling research moved from method development and property prediction to accelerated materials design and discovery guided by modelling results, data mining and machine learning [5].
- Previous studies showed that a combination of density functional theory (DFT) and machine learning (ML) methods can accelerate the discovery process for novel materials [6,7].
- In this paper, ML algorithms are explored in order to predict formation energies, E_f for LIB materials; using DFT calculated data obtained from Materials Database Project [8].

METHODOLOGY

- **Sample Construction:** The models work by performing a pre-processing step to generate a set of descriptive attributes as input features (X), using known atomic properties to generate both chemical and physical descriptors as shown in Fig.1. The true labels of the model during training (Y) are DFT calculated E_f.

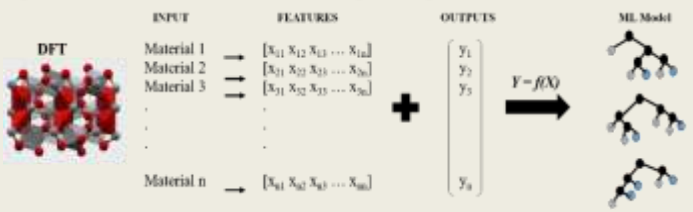

Fig. 1: From DFT to machine learning approach used to predict ΔE_f [9]

- **Model Development:** Optimised version of Catboost, Xgboost (EGB), Lightboost (LGBM), Extra Random Tree (ETR) and Random Forest (RF) models were tested for the best model selection using grid search techniques for performance boosting. ML Scikit-learn python library for supervised learning algorithm was employed.
- **Model Validation:** The accuracy of the model's predictions was evaluated by comparing DFT calculated E_f with the corresponding predicted values. Five-fold cross-validation was used to evaluate the model performance.

RESULTS AND DISCUSSIONS

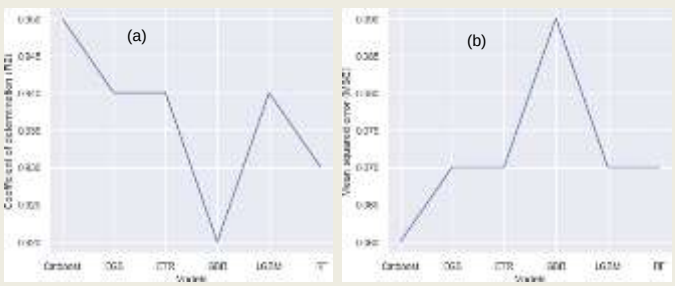

Fig. 2: (a) Determination coefficient and (b) mean square error as a function of various models

The Catboost regression model was found to be the best model with regression score and means square error of 0.95 and 0.06 eV, respectively, hence was used to predict formation energies.

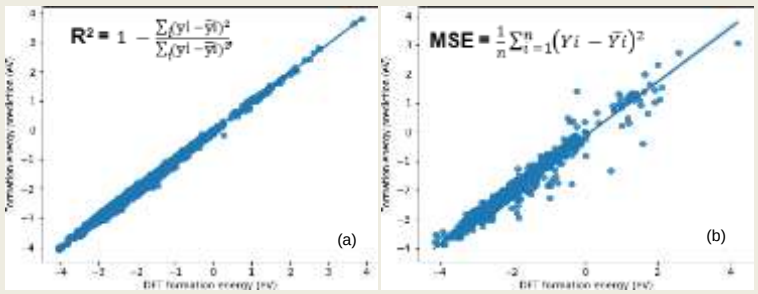

Fig. 3: Predicted formation from (a) Training and (b) test sets for the Catboost model performance

For the training data set (70 %) coefficient of determination was found to be 0.95 whereas for test set (30 %) the coefficient of determination was 0.91. Predicted values are close to 1, validating the accuracy of the Catboost model.

Table 1: Comparison of selected DFT [8] and ML formation energies

Electrodes	DFT Calculated, E_f (eV)	ML Predicted, E_f (eV)
$Li_9Mn_{12}Ni_3O_{32}$	-1.960	-1.9536
$Li_3Mn(NiO_2)_4$	-1.553	-1.4966
$Li_3MnCoNiO_6$	-1.818	-1.7599
$Li_4MnCo_2NiO_8$	-1.695	-1.7410
$LiMnO_2$	-2.171	-2.1279

The DFT calculated formation energy was obtained from Material Database Project. The ML–Catboost model's predicted formation energies agree very well with the DFT calculated within a reasonable percentage.

CONCLUSION

- Machine Learning models were successfully developed and validated, from which the formation energies of various lithium-ion battery materials were predicted.
- Amongst various algorithms that were evaluated, Catboost model was found to be the best model with the following accuracy measures:
  - coefficient of determination, $R^2 = 0.95$
  - mean square error, MSE =0.06eV.
- The results indicate that ML- Catboost model predicted DFT formation energies within a reasonable percentage error.

REFERENCES

[1] Rajendra P. Joshi, Jesse Eickholt, Liling Li, Marco Fornari, Veronica Barone, and Juan E. Peralta, ACS Applied Materials & Interfaces. 2019, 11 (20), 18494-18503

[2] Joshi R. P., Ozdemir, B.; Barone, V. and Peralta, J.E. J Phys. Chem. 2015, 6 (14), 2728-2732

[3] Bhauriyal P., Mahata, A, and Pathak, B. J. Phys. Chem C. 2017, 121 (18), 9748-9756

[4] Dunn B., Kamath, H. and Tarascon, J.M. Science. 2011, 334 (6058), 928-935

[5] Chen L. Q., Chen L. D., Kalinin S.V., Klimeck G., Kumar S.K., Neugebauer J., Terasaki I., npj Comput. Mater. 2015, 1 (15007), 1-2

[6] Hautier G., Fischer C. C., Jain A., Mueller T., Ceder G., Chem. Mater. 2010, 22 (12), 3762–3767

[7] Hautier B., Agrawal A., Kirklin S., Saal J.E., Doak J.W., Thompson A., Zhang K., Choudhary A., Wolverton C., Phys. Rev. B. 2014, 89 (9), 1-7

[8] https://materialsproject.org

[9] Maphanga R., Mokoena T., Ratsoma M., Materials Today: Proceedings. 2021, 38 (2), 773-778

ACKNOWLEDGEMENTS

NRF National Research Foundation
CSIR