Data Driven Design Rules for Transition Metal based Nano-composite Electrocatalysts for Oxygen Evolution Reaction

Vedasri Bai Khavala[a], Abhijai Velluva[a], Adhithyan Kathiravan[a], Chandan Singh Khavala[b], Harish Kuruva[a], B.S. Murty[a,c*], Tiju Thomas[a*]

[a]Department of Metallurgical and Materials Engineering, Indian Institute of Technology Madras, Chennai-600036, Tamil Nadu, India

[b]Department of Metallurgical and Materials Engineering, *Visvesvaraya National Institute of Technology, Nagpur – 440010, Maharastra, India*

[c] Indian Institute of Technology Hyderabad, Kandi -502285, Telangana, India

Email: bsm@iith.ac.in (B.S.Murty), tijuthomas@iitm.ac.in (Tiju Thomas) and vkhavala97@gmail.com (Vedasri Bai Khavala)

DATA AND CODE

To make our work reproducible, we have uploaded the following files in our GitHub repository located at https://github.com/adhi1910/TMN_OER,

1) TMON_dataset.csv – The whole dataset, consisting of 525 datapoints corresponding to 26 features, is contained in this file and was used to train the machine learning algorithms.
2) TMON_model_classifier. ipynb – This Jupyter notebook file contains the code for classification of the TMON dataset, containing all the steps of our work from the beginning to the end.
3) TMON_model_regressor. ipynb - This Jupyter notebook file contains the code for performing the regression for the TMON dataset.
4) TMO_data.csv - The whole dataset, consisting of 295 TMO datapoints corresponding to 25 features, is contained in this file and was used to train the machine learning algorithms

S1. LIST OF FEATURES

The categorical (11) and continuous (15) features in TMON dataset. The TMO dataset contains same features excluding the oxided or nitrided categorical feature i.e., TMO dataset contains 10 categorical and 15 continuous features. The following features being used in the model highlight both materials and device perspective and is indicated in brackets following the feature:

- **Categorical features**

The encoding used for the corresponding feature are indicated using single quotes ‘’:

1. Compound Class (materials perspective):

(i) cobalt class (‘Co’), (ii) nickel class (‘Ni’), (iii) iron class (‘Fe’), and (iv) other transition metals class (molybdenum, vanadium, copper, manganese, chromium, zinc & titanium) (‘Other TMs’).

2. Composite (Y or N) (materials perspective):

Composite= ‘1’, non-composite = ‘0’

3. Oxided or nitrided (materials perspective):

'1' if nitrided; otherwise, '0'

4. Presence of Carbon (materials perspective):

Yes=’1’, No =’0’

5. Carbon Structures (materials perspective):

(i) no carbon ‘0’, (ii) 1D carbon structures (carbon nanotube, carbon nanowires, carbon fibers, etc.) ‘1’, (iii) 2D carbon structures (graphene, carbon nanosheets, etc.) ‘2’, (iv) 3D

carbon structures (carbon cloth, carbon nanocube, crumbled graphene, etc.) ‘3’ and (v) reduced graphene oxide ‘4’

6. N-doped Carbon Structure (0 or 1) (materials perspective)

   ‘0’ is absence of nitrogen, ‘1’ presence of nitrogen

7. Doping (other than TMs) (materials perspective):

   ‘0’ means undoped, ‘1’ means doped

8. Morphology Category (materials perspective):

   (i) bulk ‘1’ (ii) nanorod, nanotubes, nanowires, nanofibers, nanoneedles, etc. ‘2’ (iii) nanoparticles, nanospheres & microspheres ‘3’ (iv) combinatorial structures (such as core-shell, a structure with multiple morphologies) ‘4’ (v) thin films ‘5’ (vi) 3D nanomaterials with unique morphology (flower-like, urchins, nanocrystals, nano boxes, nanocubes, etc.) ‘6’ (vii) nanosheets, nanomeshes, nanoflakes, nanoplates, etc. ‘7’ (viii) powders ‘8’.

9. Grown on (materials perspective):

   This feature indicates the substrate on which the catalyst is synthesized on.

   (i) free-standing/ not grown on any substrate (‘0’), (ii) nickel foam or nickel foil (‘1’), (iii) carbon matrixes (‘2’), and (iv) Other metallic foams, ligaments, nanosheets and foils (‘3’).

10. Porous - 1, Non-Porous – 0 (materials perspective)

11. Substrates (device perspective):

    This feature indicates the substrate on which electrochemical measurements are performed for the catalyst.

nickel foam (‘NF’), glassy carbon (‘GC’), carbon-based (other than GC, like carbon cloth (CC), carbon paper (CP), carbon fiber paper (CFP), graphite disk, etc.) (‘carbon_based’), and TM based (iron foam, copper foam, titanium foil, titanium mesh, etc.) (‘TM_based’).

- **Continuous features**

12. Presence of atoms (other than TMs) (materials perspective):

    filled by the atomic number of respective atoms

13. Cobalt atoms (materials perspective):

    Filled with number of atoms

14. Nickel atoms (materials perspective):

    Filled with number of atoms

15. Iron atoms (materials perspective):

    Filled with number of atoms

16. Molybdenum atoms (materials perspective):

    Filled with number of atoms

17. Manganese atoms (materials perspective):

    Filled with number of atoms

18. Copper atoms (materials perspective)

    Filled with number of atoms

19. Zinc atoms (materials perspective):

    Filled with number of atoms

20. Chromium Atoms (materials perspective):

    Filled with number of atoms

21. Vanadium Atoms (materials perspective):

Filled with number of atoms

22. Tungsten atoms (materials perspective):

Filled with number of atoms

23. Titanium atoms (materials perspective):

Filled with number of atoms

24. Average d electrons (materials perspective):

Filled by taking average of total number of d-electrons available in the electrocatalyst.

$Avg.d-electrons = \frac{total\ no.of\ d-electrons}{\text{total no of atoms}}$, for non-composite

$$Avg.d-electrons = \frac{avg.no\ of\ d-electrons\ in\ phase\ 1 + avg.no\ of\ d-\ electrons\ in\ phase\ 2 + \ldots avg.no\ of\ d-electrons\ in\ phase\ N)}{total\ number\ of\ phases\ (N)}$$

, for composite

25. pH (device perspective):

pH values of the respective electrolyte used

26. Catalyst Loading (mg/cm2) (device perspective)

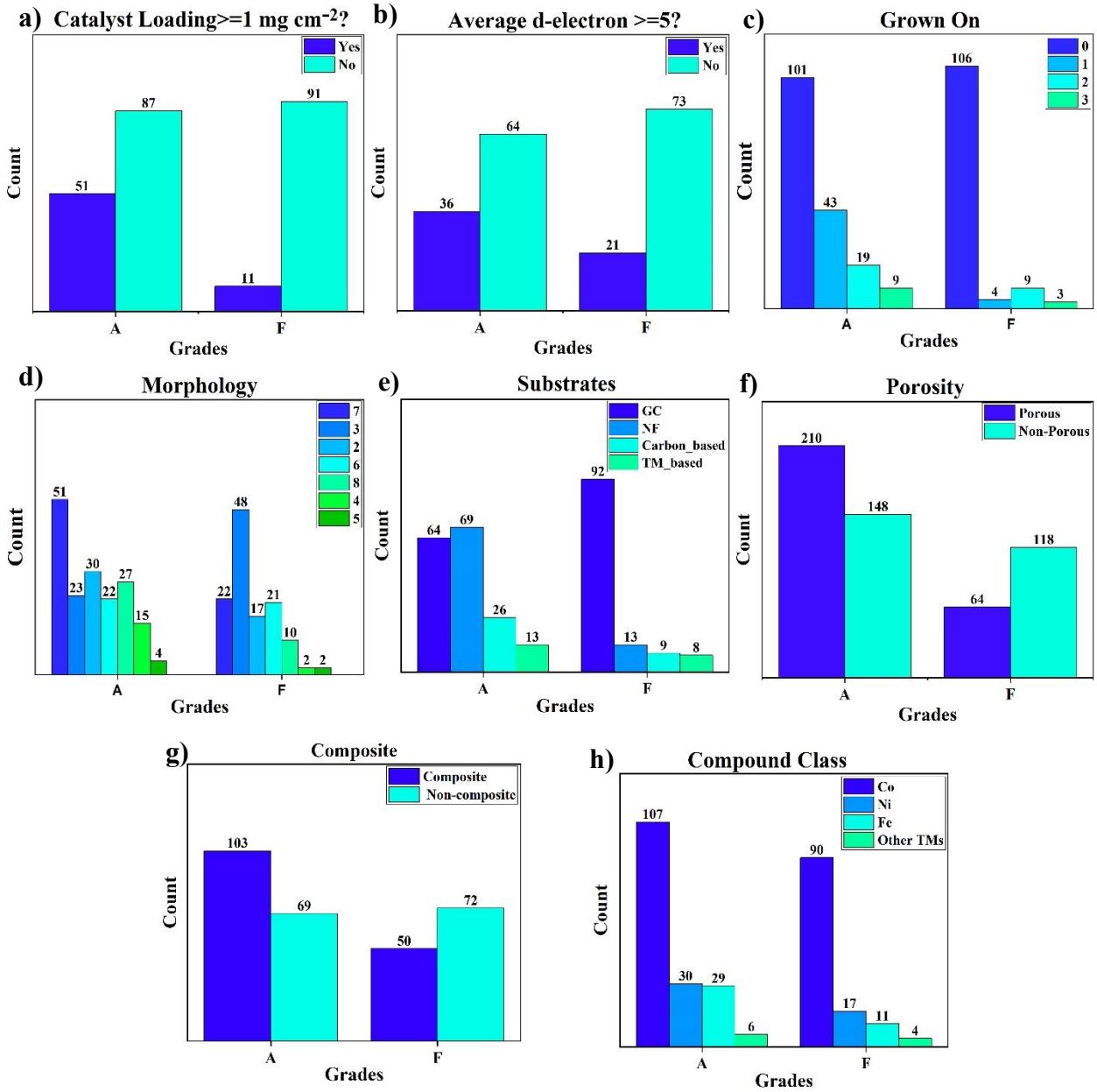

Figure S1: TMO dataset representation (a) Catalyst loading, (b) Average d-electrons, (c) Grown on, (d) Morphology, (e) Substrates, (f) Porosity, (g) Composite, and (h) Compound class

Table S1: Probability of Compound Class, Composite, Morphology types, Grown On, Porous, substrates, catalyst loading >=1 and average d-electrons >=5 sub-classes being grade 'A' for TMON dataset. The category which is highlighted in green will have more probability of being grade 'A', compared to other categories in respective feature.

Feature	Categories	Probability of Being Grade 'A' (in %)
Compound Class	Co	62.67
	Ni	**73.68**
	Fe	**69.12**
	Other TMs	53.33
Composite	**Yes**	**74.19**
	No	53.92
Morphology Category	1	54.54
	2	75.47
	3	45.97
	4	**75.75**
	5	75
	6	66.67
	7	**72.65**
	8	**70.45**
Grown On	0	56.85
	1	91.67
	2	76.74
	3	72.72
Porous	0	49.38
	1	**69.4**
Substrates	GC	41.02
	NF	84.15
	Carbon_based	74.28
	TM_Based	61.9
Catalyst loading >=1	**Yes**	**87.07**
	No	64.21
Average d-electrons >=5	**Yes**	**74.81**
	No	56.22
Oxided/Nitrided	**Nitrided**	**75.11**
	Oxided	58.5

Table S2: Probability of Compound Class, Composite, Morphology types, Grown On, Porous, substrates, catalyst loading >=1 and average d-electrons >=5 sub-classes being grade 'A' for TMO dataset. The category which is highlighted in green will have more probability of being grade 'A', compared to other categories in respective feature.

Feature	Categories	Probability of Being Grade 'A' (in %)
Compound Class	Co	54.31
	Ni	**63.83**
	Fe	**72.5**
	Other TMs	60
Composite	**Yes**	**67.32**
	No	48.93
Morphology Category	2	63.82
	3	32.39
	4	**88.24**
	5	66.67
	6	51.16
	7	**69.86**
	8	**72.97**
Grown On	0	48.79
	1	91.48
	2	67.85
	3	75
Porous	0	49.38
	1	**69.4**
Substrates	GC	41.02
	NF	84.15
	Carbon_based	74.28
	TM_Based	61.9
Catalyst loading >=1	**Yes**	**73.27%**
	No	48.87
Average d-electrons >=5	**Yes**	**63.15**
	No	57.38

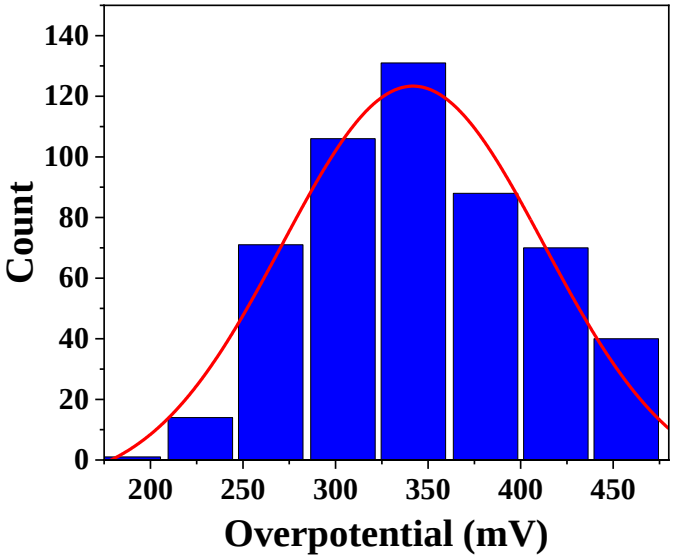

Figure S2: Distribution of the target variable η_{10} in TMON dataset, which has a range of 180 to 460 mV and a mean of 341.85 mV.

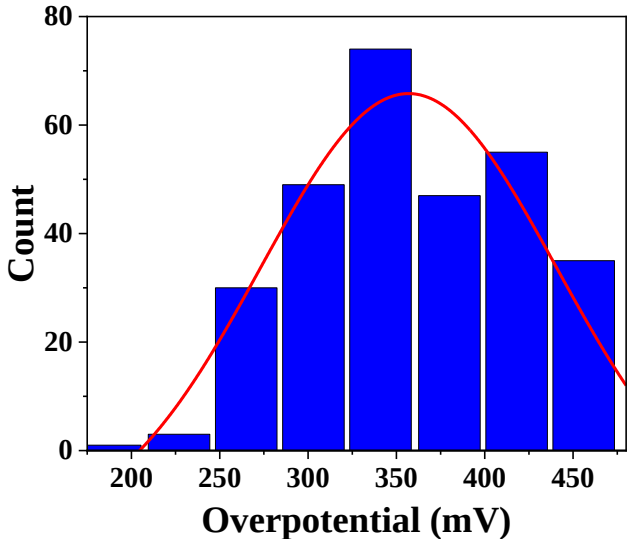

Figure S3: Distribution of the target variable η_{10} in TMO dataset. The distribution of is in the range of 180 to 460 mV with a mean of 356.9 mV.

S2. FEATURE ENGINEERING AND DATA PRE-PROCESSING

Each compound is introduced to the model by multiplying the number of transition metal atoms present in the compound by their respective atomic masses. The "average d-electrons" feature is calculated for all possible data points, which is an important feature that tunes the catalytic properties. The missing values in "catalyst loading" and "average d-electrons" are filled with the mean of the distribution. The nominal categorical features such as "compound class" and "substrates" are label encoded. Figure S4 for TMONs and Figure S5 for TMOs represent the heat map correlation matrix which gives Pearson correlation (-1: perfect negative correlation, 0: no correlation, and +1: perfect positive correlation) between the various features used in the model and found that no correlation between the continuous variables in the final dataset was higher than 0.75).

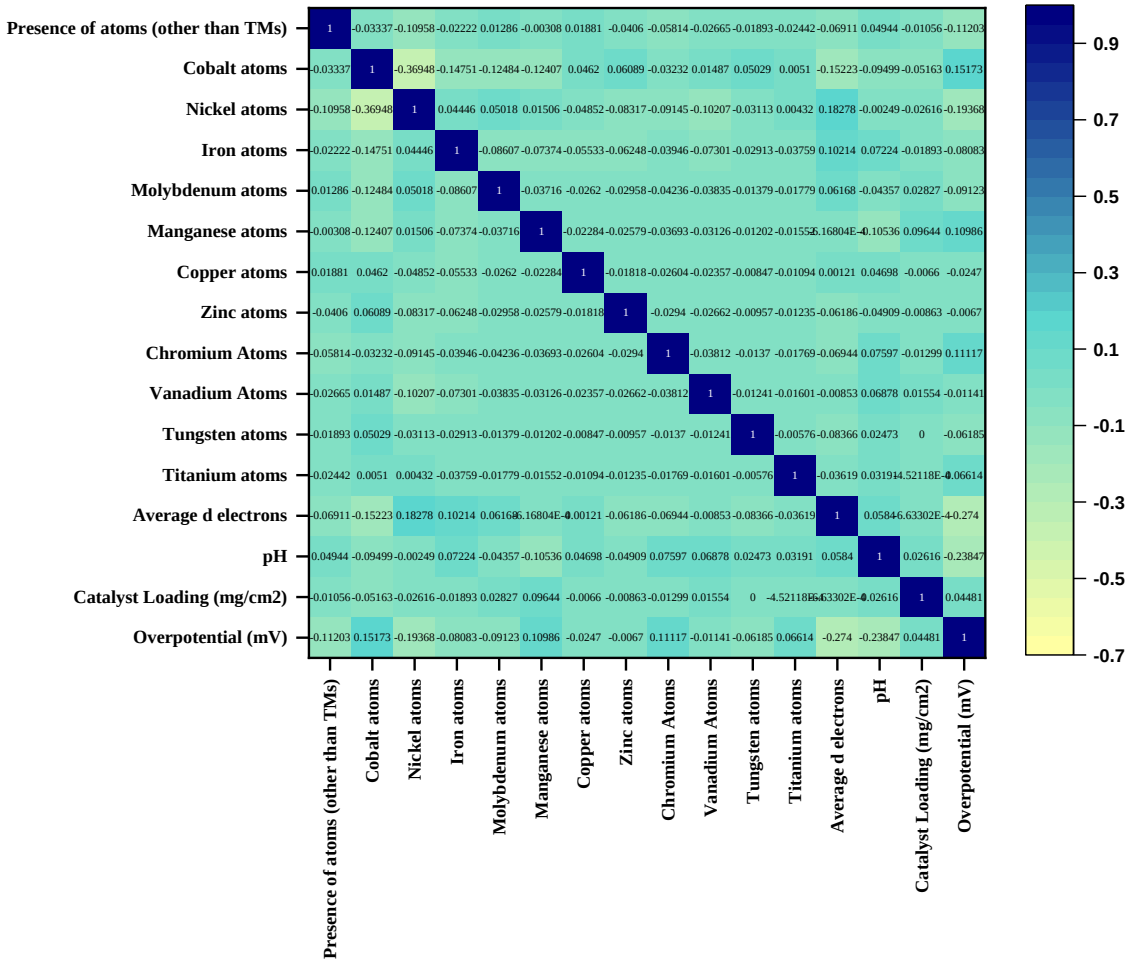

Figure S4: Heat map showing the Pearson correlation between features. where a value of -1 denotes a perfect negative correlation, a value of 0 denotes no correlation at all, and a value of +1 denotes a perfect positive correlation. No correlation between the continuous variables in the TMON dataset exceeds 0.75.

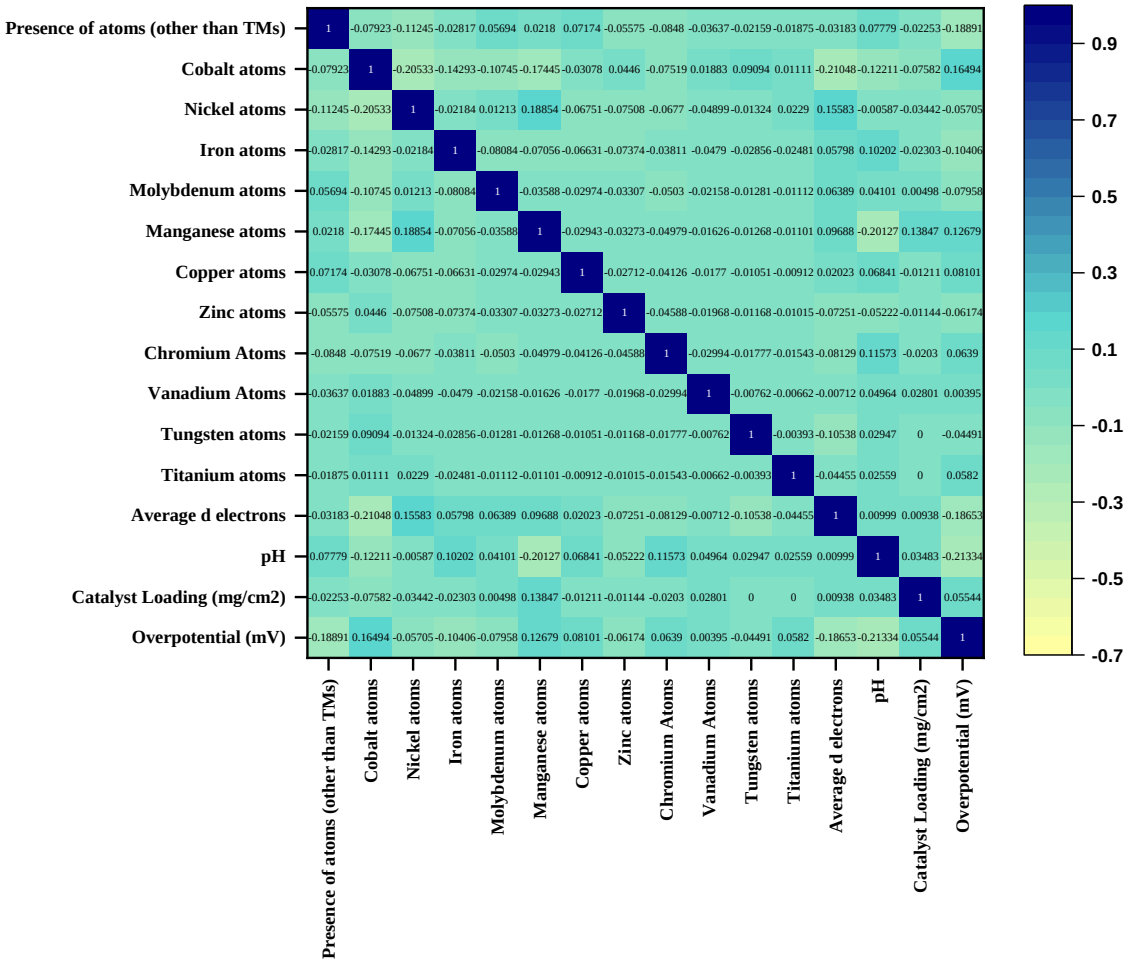

Figure S5: Heat map showing the Pearson correlation between features. where a value of -1 denotes a perfect negative correlation, a value of 0 denotes no correlation at all, and a value of +1 denotes a perfect positive correlation. No correlation between the continuous variables in the TMO dataset exceeds 0.75.

S3. PARAMETERS TO VALIDATE THE MODELS

The parameters to validate the ML classification models are precision, recall, F1 score, and accuracy. All these parameters are related to the confusion matrix, which summarizes the performance of the classification model. The confusion matrix will be an N x N matrix for a classification model of' N' classes. The error used to validate the ML regression model is root mean square error (RMSE) and mean absolute percentage error (MAPE).

S3.1. Precision

Precision is the ratio of true positives to the sum of true positives and false positives. As the name suggests, the Precision (P) metric represents the ability of an ML model not to label a data positive that is negative. If the ML model does not produce false positives, its P value will be 1.

S3.2. Recall

Recall (R or sensitivity) refers to the total relevant result correctly classified by the ML model. In each class, recall is defined as the ratio of true positives to the sum of true positives and false negatives [50]. The higher the value of R, a model can classify the data in its actual classes.

S3.3. F1 Score

The F- measure (F1 score) is a widely used metric to analyse the performance of a classification model when the target classes are imbalanced. The F1 score is the harmonic mean of precision and recall, equal weightage to both [51].

$$F1\ score = \frac{2}{\frac{1}{P} + \frac{1}{R}} = \frac{2PR}{P + R}$$

S3.4. Accuracy

The accuracy score is the ratio of correct predictions to the total number of predictions. In the case of a classification model, it will be the ratio of the sum of true positives and true negatives to the total number of datasets. In imbalanced classifications, the F1 score is more useful than the accuracy score.

S3.5. RMSE

Root Mean Square Error (RMSE) represents the standard deviation of residuals (prediction errors). RMSE gives relatively high weightage to larger errors and indicates the absolute fit of the model. RMSE is defined as;

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (predicted_i - actual_i)^2}{n}}$$

S3.6. MAPE

Mean Absolute Percentage Error (MAPE) is a relative error measure that uses absolute values to keep the positive and negative errors from cancelling one another out and uses relative errors. MAPE is defined as;

$$MAPE = \frac{1}{n}\sqrt{\frac{\sum_{i=1}^{n} Actual_i - Predicted_i}{Actual_i}}$$

S4. FEATURE IMPORTANCE

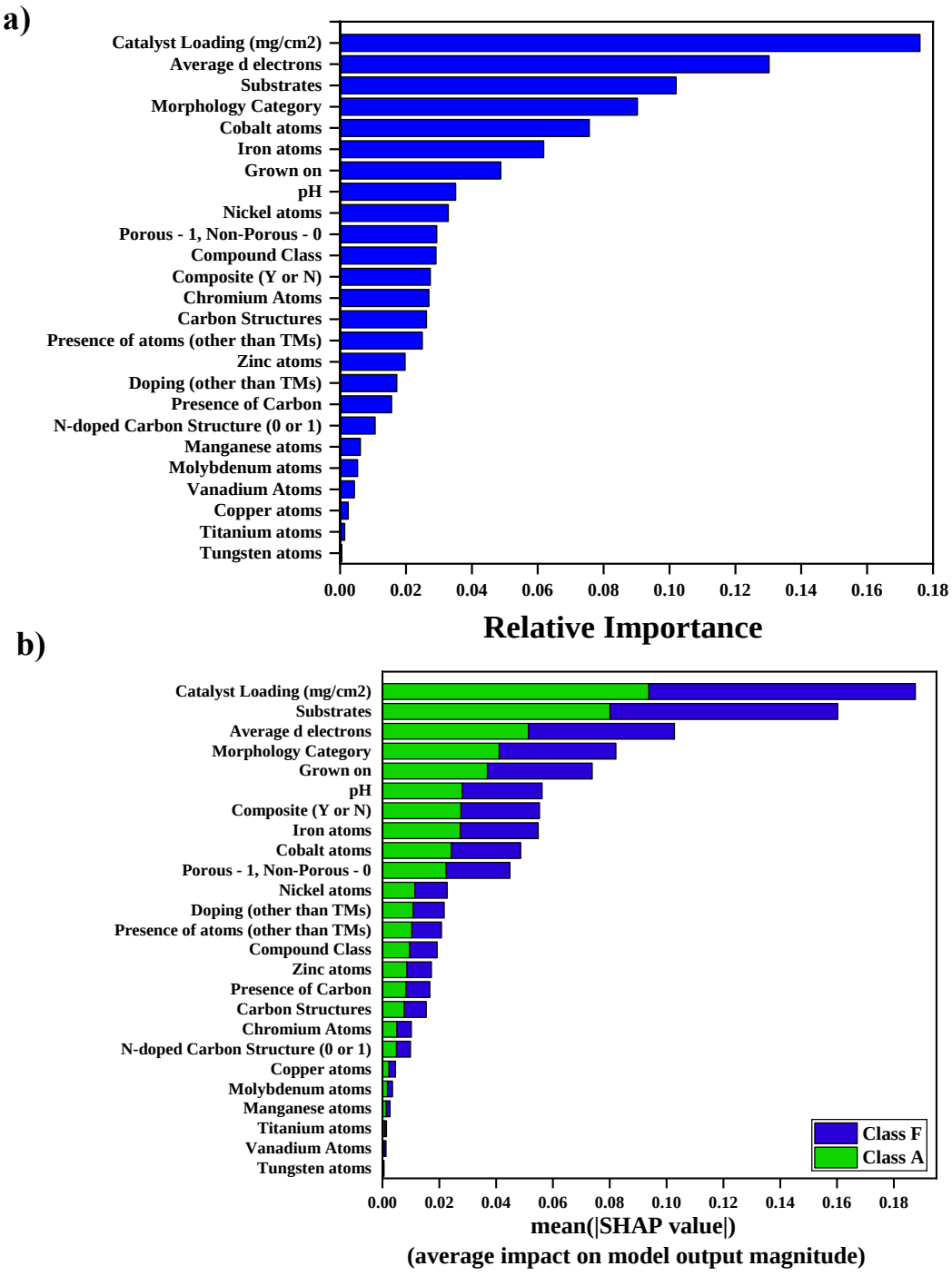

Figure S6: (a) Feature Importance Plots of trained features of RF model trained on the collected TMO dataset, with the catalyst loading and average d-electrons as the highest priority. (b) SHAP analysis of trained features of RF model trained on TMO dataset showing the most important features are catalyst loading, substrates, average d electrons and morphology category.

Neglecting the order of features, the number of discrepancies between FRF and SHAP is only one among first five, first ten as well as first fifteen in TMONs (Figure 2 (a) and (b)). Similar comparison between FRF and SHAP for TMO dataset is given in Figure S6(a) and (b).In TMO dataset the number of discrepancy between FRF and SHAP features are one among first five (grown on in SHAP is replaced by cobalt atoms in FRF), one among first ten (composite (Y or N) in SHAP been replaced by Nickel atoms), and one among first fifteen (zinc atoms and doping (other than TMs)in SHAP is replaced by chromium atoms and carbon structures).

SHAP values capture the average contribution of each feature to the prediction across all possible feature subsets. This approach provides a more nuanced and consistent measure of feature importance, as it accounts for interactions between features. Traditional methods (e.g., Random Forest, Gradient Boosting) might not capture interactions between features effectively. They can underestimate the importance of features when their impact is context-dependent and varies depending on the presence of other features. SHAP values naturally account for interactions between features. They consider how the inclusion or exclusion of a feature affects the overall prediction, capturing both main effects and interactions.

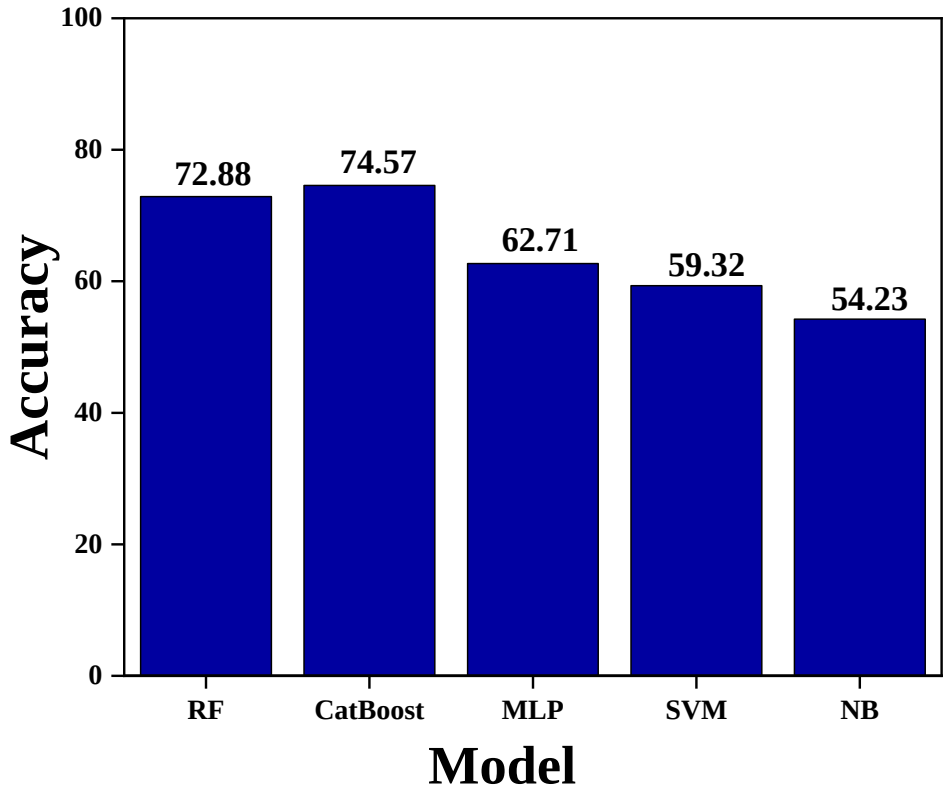

Figure S7. Comparison of accuracy of various ML models trained on TMO dataset. Using TMO dataset the most accurate RF model shows F1 scores of 0.82 and 0.73 for grades 'A' and' F,' respectively

Table S3: Values of optimized hyperparameters for each ML algorithm tested.

Algorithm	**Values of optimized hyperparameters**
RandomForestClassifier	class_weight='balanced', max_features=log2, n_estimators=100, random_state=0
CatBoostClassifier	iterations=500, depth=5, random_state=0
SupportVectorClassifier	Gamma = scale, c = 1.0, random_state=0
GaussianNaiveBayesClassifier	var_smoothing: 1e-09
MultiLayerPerceptron classifier	activation: relu, alpha: 0.000, momentum': 0.9, solver: adam, random_state=0

Table S4: Comparison of different ML models during the testing phase of overpotential grade prediction for the TMON dataset

	RF	**CatBoost**	**SVM**	**MLP**	**NB**
Accuracy (%)	77.14	76.19	67.61	63.8	40.95
F1 score [A, F]	[0.84, 0.61]	[0.84, 0.56]	[0.80, 0.11]	[0.74, 0.41]	[0.74, 0.34]
Precision [A, F]	[0.78, 0.73]	[0.76, 0.76]	[0.67, 1.00]	[0.70, 0.46]	[0.73, 0.36]
Recall [A, F]	[0.90, 0.53]	[0.93, 0.44]	[1.00, 0.06]	[0.78, 0.36]	[0.89, 0.24]

Table S5: Comparison of different ML models during the testing phase of overpotential grade prediction for the TMO dataset.

	RF	**CatBoost**	**SVM**	**MLP**	**NB**
Accuracy (%)	72.88	74.57	59.32	62.71	54.23
F1 score [A, F]	[0.78, 0.65]	[0.79, 0.67]	[0.74, 0.00]	[0.66, 0.59]	[0.63, 0.40]
Precision [A, F]	[0.76, 0.68]	[0.76, 0.71]	[0.59, 0.00]	[0.72, 0.53]	[0.61, 0.43]
Recall [A, F]	[0.80, 0.62]	[0.83, 0.62]	[1.00, 0.00]	[0.60, 0.67]	[0.66, 0.38]

S5. RF regressor with TMON dataset

Table S6: Values of optimized parameters for RF Regression on TMON dataset

Parameter	Value
'bootstrap'	'True'
'ccp_alpha'	0.0
'criterion'	'squared_error'
'max_depth'	'None'
'max_features'	'sqrt'
'max_leaf_nodes'	'None'
'max_samples'	'None'
'min_impurity_decrease'	0.0
'min_samples_leaf'	1
'min_samples_split'	2
'min_weight_fraction_leaf'	0.0
'n_estimators'	100
'n_jobs'	'None'
'oob_score'	'False'
'random_state'	0
'verbose'	0
'warm_start'	'False'

Errors:

Table S7: Values of the error metrics for RF Regression on TMON dataset

Errors	Error Value (mV)
MSE	2043.293
RMSE	45.203
MAE	36.468
MAPE	0.115

S6. RF regressor with TMO dataset

Table S8: Values of optimized parameters for RF Regression on TMO dataset

Parameter	Value
'bootstrap'	'True'
'ccp_alpha'	0.0
'criterion'	'squared_error'
'max_depth'	'None'

'max_features'	'sqrt'
'max_leaf_nodes'	‘None’
'max_samples'	‘None’
'min_impurity_decrease'	0.0
'min_samples_leaf'	1
'min_samples_split'	2
'min_weight_fraction_leaf'	0.0
'n_estimators'	100
'n_jobs'	‘None’
'oob_score'	‘False’
'random_state'	0
'verbose'	0
'warm_start'	‘False’

Errors:

Table S9: Values of the error metrics for RF Regression on TMO dataset

Errors	**Error Value (mV)**
MSE	1803.069
RMSE	42.463
MAE	32.691
MAPE	0.095