

Article

# Machine Learning Prediction of Surface Segregation Energies on Low Index Bimetallic Surfaces

Damilola Ologunagba and Shyam Kattel \*

Department of Physics, Florida A&M University, Tallahassee, FL 32307, USA; damilola1.ologunagba@famou.edu

\* Correspondence: shyam.kattel@famou.edu

Received: 1 April 2020; Accepted: 24 April 2020; Published: 1 May 2020



**Abstract:** Surface chemical composition of bimetallic catalysts can differ from the bulk composition because of the segregation of the alloy components. Thus, it is very useful to know how the different components are arranged on the surface of catalysts to gain a fundamental understanding of the catalysis occurring on bimetallic surfaces. First-principles density functional theory (DFT) calculations can provide deeper insight into the surface segregation behavior and help understand the surface composition on bimetallic surfaces. However, the DFT calculations are computationally demanding and require large computing platforms. In this regard, statistical/machine learning methods provide a quick and alternative approach to study materials properties. Here, we trained previously reported surface segregation energies on low index surfaces of bimetallic catalysts using various linear and non-linear statistical methods to find a correlation between surface segregation energies and elemental properties. The results revealed that the surface segregation energies on low index bimetallic surfaces can be predicted using fundamental elemental properties.

**Keywords:** catalysts; surface segregation energy; density functional theory; machine learning; bimetallic alloys

## 1. Introduction

Transition metals (TM) catalyze many catalytic and electrocatalytic reactions [1–7]. The catalytic activity and selectivity of transition metal-based catalysts can be altered significantly by adding a second/guest metal in host metals [6,8–12]. The addition of the second/guest metal in host metal could result in the formation of an ordered bimetallic alloy with a well-defined crystal structure. It has been shown over the years for many reactions that bimetallic alloy catalysts show superior activity compared to monometallic catalysts [6,12–14]. For example, oxygen reduction reaction (ORR) activity of PtNi, PtCo and PtFe bimetallic alloy catalysts is superior to monometallic Pt catalyst [11,13–15]. Thus, alloying has been found to be an effective way to improve the activity of TM based catalysts in heterogeneous catalysis [6,12,16–18].

The surface composition of catalysts is of paramount importance in catalysis as the reaction centers/active sites are located on the surface where reactions occur allowing the bond forming and breaking processes. For a bimetallic catalyst, its surface chemical composition can differ from the bulk composition because of the segregation of the alloy components [19]. As a result, the segregation could result in the surface composition of an alloy catalyst enriched in one component. For instance, the surface of PtTM alloy (TM = 3d-TM) catalysts is enriched with Pt forming a core-shell type of structures due to the preferred segregation of Pt on the surface [3,20–25]. Therefore, the understanding of the surface composition of TM-alloy catalysts provides important insight for the rational design of alloy catalysts. Experimental techniques such as photoemission spectroscopy of surface core-level shifts (SCLS) and quantum computational methods, such as first-principles density functional theory (DFT) calculations, have been routinely used to study the surface composition and segregation trends

on alloy catalysts [26–30]. In many alloy surfaces, the difference in surface energy between the metals that make up the alloy and the heat of solution of the guest metal into the bulk of the host metal is found to be the driving forces of surface segregation [31]. The metal with lower surface energy and or positive heat of solution has been found to move to the surface.

Ruban et al. [32] have performed calculations based on Green's-function linear-muffin-tin-orbitals method to estimate the surface segregation energy on the thermodynamically most stable low index surfaces [(111) surface of face centered cubic phase, (110) surface of body centered cubic phase and (001) surface of hexagonal cubic phase] of bimetallic alloys. Similarly, DFT calculations were carried out to understand segregation behavior on low-index open surfaces [(100) surface of fcc and bcc and (1010) surface of hcp] of bimetallic transition metal alloys [33]. The DFT calculations provide an accurate prediction of surface segregation trends and are often consistent with the experimental observations [30,33–36]. However, the DFT calculations are computationally demanding and require large computing platforms. In this regard, statistical/machine learning methods provide a quick and alternative approach to study materials properties. In recent years, such models have been successfully used to study the structures of catalysts as well as reaction mechanisms in heterogeneous catalysis [37–45]. Here, we used a previously published database of surface segregation energies calculated using Green's-function method [32] to fit various machine learning models. The surface segregation energies were correlated with elemental properties of host and guest metals. Results obtained from the various statistical models used in this study demonstrated that the surface segregation energies of bimetallic alloys can be predicted using simple non-linear regression methods.

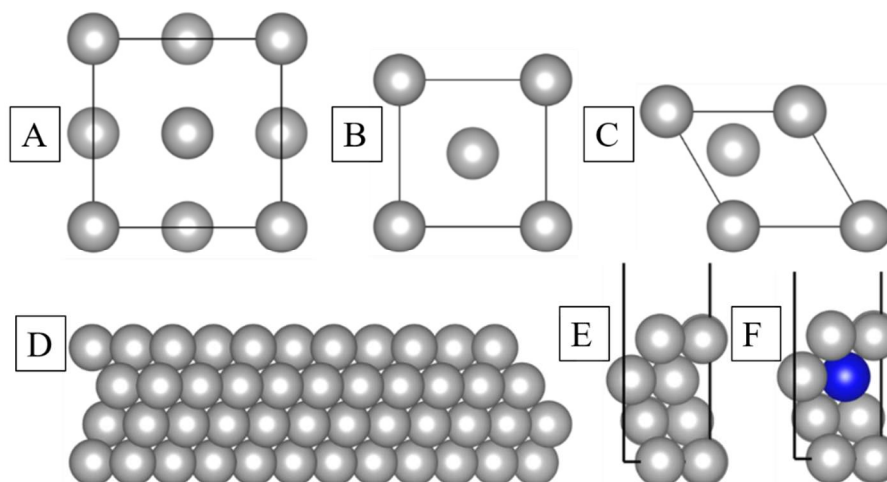
## 2. Methods

The dataset used in the present study has one response variable which is the surface segregation energies of transition metal alloys and a total of 20 features (10 features each for host and guest metals) as shown in Table 1 and Table S7 (in the supporting information (SI)). Surface segregation energies were taken from the study of Ruban et al. [32]. The features were the elemental properties of the host and guest metals (Table 1) [46]. These properties are atomic radius, atomic number, atomic mass, density, work function, period, electron affinity, ionization energy, d-shell and Pauling electronegativity. Each of the host metals was combined with all the guest metals to form a bimetallic system. With each of the 24 metals (Ti, V, Cr, Mn, Fe, Co, Ni, Cu, Zr, Nb, Mo, Tc, Ru, Rh, Pd, Ag, Hf, Ta, W, Re, Os, Ir, Pt, Au) combined with all the 24 metals, we had a total of ~576 systems. The host transition metals were classified according to three most common crystal structures—face centered cubic (FCC), body centered cubic (BCC) and the hexagonal closed packed (HCP) [47] as shown in Figure 1A–C, respectively. The surface segregation energy was defined as the energy required to move an impurity atom from the inside of a host crystal to the surface (Figure 1D,F) [33]. The boxplots (Figure S25) showed that that two features (density and atomic mass) showed relatively large inhomogeneity compared to other features included in this study when compared between host and guest atoms.

Various linear and non-linear machine learning models were employed to correlate the elemental properties of host and guest metals with surface segregation energies of bimetallic alloys. Linear regression is a method that finds the linear relationship between descriptors and one or more target variables. The linear models used in the present study include Ordinary Least Squares (OLS) and Partial Least Squares (PLS) [48]. The non-linear regression is a method that models the non-linear relationship between the descriptors and the target variable. The non-linear models used in this study included gradient boosting regression (GBR) [49], Gaussian process regression (GPR) [50], kernel ridge regression (KRR) [51] and support vector regression (SVR) [52] (see Figures S1–S20 in SI) for all the models except GBR).

**Table 1.** Elemental properties of transition metals used as features.

Metals	Radius/pm	Atomic No.	Atomic Mass/amu	Period	Ionization Energy/KJ/mol	Density/Kg/m <sup>3</sup>	Electron Affinity/eV	Work Function/eV	Pauling Electronegativity	d-Shell
Ti	140	22	47.87	4	658.81	4507	0.080	4.33	1.54	3
V	135	23	50.94	4	650.91	6110	0.526	4.30	1.63	3
Cr	140	24	51.99	4	652.87	7140	0.667	4.50	1.66	3
Mn	140	25	50.94	4	717.28	7470	0.005	4.10	1.55	3
Fe	140	26	55.85	4	746.47	7874	0.164	4.50	1.83	3
Co	135	27	58.93	4	760.40	8900	0.662	5.00	1.88	3
Ni	135	28	58.69	4	737.14	8908	1.157	5.15	1.91	3
Cu	135	29	63.55	4	745.48	8920	1.228	4.51	1.90	3
Zr	135	40	91.22	5	640.10	6511	0.427	4.05	1.33	4
Nb	145	41	92.91	5	652.13	8570	0.894	4.30	1.60	4
Mo	145	42	95.95	5	684.32	10,280	0.747	4.60	2.16	4
Tc	135	43	98.00	5	686.92	11,500	0.550	4.90	1.90	4
Ru	130	44	101.07	5	710.18	12,370	1.050	4.71	2.20	4
Rh	135	45	102.91	5	719.67	12,450	1.138	4.98	2.28	4
Pd	140	46	106.42	5	804.38	12,023	0.558	5.12	2.20	4
Ag	160	47	107.87	5	731.00	10,490	1.303	4.26	1.93	4
Hf	155	72	178.49	6	658.52	13,310	0.100	3.90	1.30	5
Ta	145	73	180.95	6	728.42	16,650	0.323	4.25	1.50	5
W	135	74	183.84	6	758.76	19,250	0.816	4.55	2.36	5
Re	135	75	186.21	6	755.82	21,020	0.120	4.87	1.90	5
Os	130	76	190.23	6	814.17	22,610	1.120	4.83	2.20	5
Ir	135	77	192.22	6	865.19	22,650	1.566	5.27	2.20	5
Pt	135	78	195.08	6	864.40	21,090	2.128	5.65	2.28	5
Au	135	79	196.97	6	890.13	19,300	2.309	5.10	2.54	5



**Figure 1.** Unit cells of bulk (A) face centered cubic (FCC), (B) body centered cubic (BCC) and (C) hexagonal closed packed (HCP) crystal structures, (D) a low index (111) surface of FCC crystal, (E) a four layer  $2 \times 2$  unit cell of (111) surface of FCC crystal and (F) a four layer  $2 \times 2$  unit cell of (111) surface of a bimetallic alloy. Grey: host atom and dark blue: guest atom.

All the models were implemented using an open source scikit-learn library of python [53]. Non-linear models were tested for a range of values for various parameters. The parameters for GBR were the number of estimators, learning rate and max depth, the parameters for GPR were kernel and the number of restart optimizers, the parameters for KRR were alpha, coefficient, degree, gamma, kernel and kernel parameters. The parameters that produced the best model predictions for the training set were used for the prediction of the surface segregation of the test set. The accuracies of our models were primarily tested using the root mean square error (RMSE) and mean absolute error (MAE), commonly used metrics in machine learning/statistical models.

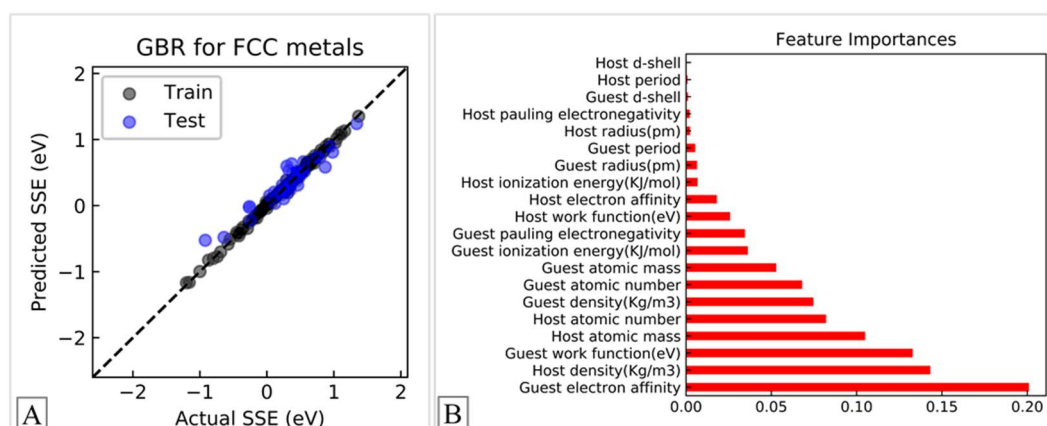
For the GBR method, we tested different values of parameters, including the number of estimators, the learning rate and the maximum depth. We tested the learning rate values of 0.011, 0.1 and 0.2. For the number of estimators, we tested values ranging from 100 to 300. Also, we tested maximum depth values ranging between 3 and 5. Based on the results, we found that for bimetallic FCC (111) surfaces, the number of estimators, learning rate and the maximum depth values that produced the best accuracy were 200, 0.1 and 3, respectively. For BCC (110) surfaces, the number of estimators, learning rate and the maximum depth values that produced the best accuracy were 100, 0.1 and 3, respectively. Similarly, for HCP (001) surfaces, the number of estimators, learning rate and the maximum depth values that produced the best accuracy were 100, 0.1 and 3, respectively. For combined FCC (111), BCC (110) and HCP (001) dataset, the number of estimators, learning rate and the maximum depth values that produced the best accuracy were 200, 0.1 and 3, respectively.

We employed the random permutation cross-validation, also known as the shuffle split cross-validation technique, to offset the bias of our model. This method was similar to the k-fold cross-validation where the data is split into k-folds and one of the folds will be the test set and the other  $(k - 1)$  folds will be the training set. The difference between the k-fold cross-validation and the shuffle split cross-validation was that for the former, the data was not shuffled after each iteration. However, for the latter, the data was shuffled and then split into training and test set. The data set was shuffled and split 7 times and randomly sampled during each iteration. The predictions using the test and training sets in each iteration is not used in the development of the model. Rather, all the predicted values using the test and training sets in each iteration were appended to a list and the error is calculated over the entire dataset. In the case of the shuffle split cross-validation method, splitting our data into a test set (20%) and a training set (80%) resulted in the best accuracy. (see SI for the details of our results using the shuffle split cross-validation).

### 3. Results and Discussion

#### 3.1. Model Predictions of the Surface Segregation Energy on Bimetallic FCC (111) Surfaces

The surface segregation energies on bimetallic FCC (111) surfaces were modeled using OLS, PLS, SVR, GPR, GBR and KRR methods. It was found that the GBR model showed the best result (Figure 2A) with an RMSE of 0.15 eV for the test set and 0.03 eV for the training set. Our calculated MAE of training and test sets were 0.02 eV and 0.11 eV, respectively. As shown in Table 2, the  $R^2$  values for the training and test sets were 0.99 and 0.89, respectively. In contrast, the RMSE and MAE values obtained using other machine learning models (Figures S1–S5; Tables S1–S6) were higher compared to those obtained using the GBR method. The GBR method typically promotes weak learners to strong learners and learn complex non-linear decision boundaries via boosting to increase its overall predictive performance. As expected, the predictive performance of the GBR method, as well as the other methods, was found to be much better for the training set as compared to the test set. Among all the data points, the predicted value of surface segregation energy of RhAg (−0.52 eV) had the largest deviation from actual value (−0.92 eV). Importantly, our GBR model was able to capture the sign of surface segregation energies on (111) surface of all bimetallic alloys and the RMSE of the GBR model was comparable to the typical error (~0.1 eV) in density functional theory (DFT) calculations. For example, consistent with the experimental observations, our model predicted Pt surface segregation on (111) surfaces of PtFe, PtCo and PtNi bimetallic systems [11,15]. Thus, the results based on our small data set showed that the GBR model can be used to reliably predict surface segregation energies on thermodynamically most stable low index bimetallic (111) surfaces using the elemental properties (20 features) of components in bimetallic alloys.



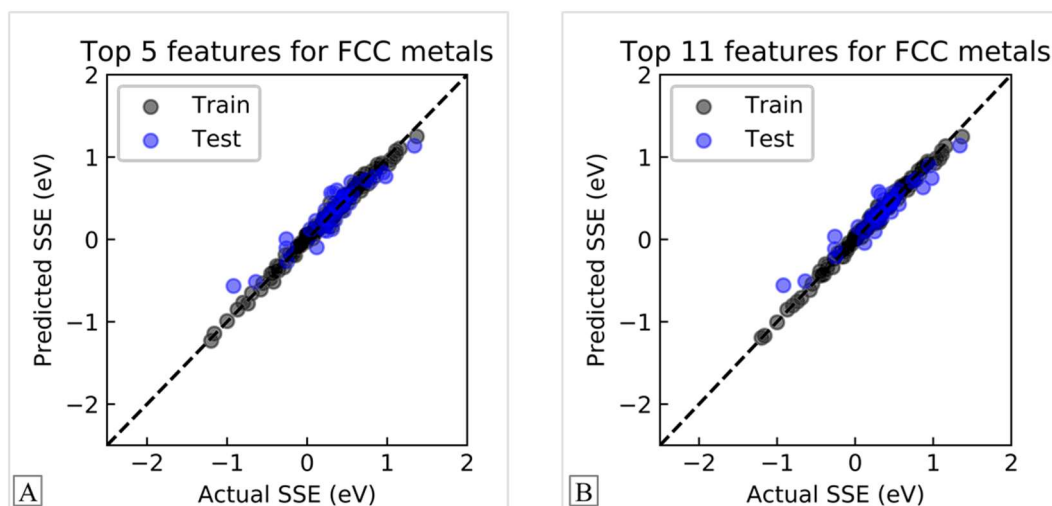
**Figure 2.** (A) Model predictions and (B) Feature importance scores for FCC(111) surfaces obtained from GBR model. SSE = surface segregation energy.

**Table 2.** Root mean square error (RMSE), mean absolute error (MAE) and  $R^2$  values for training and test sets obtained from gradient boosting regression (GBR) model using 20 features.

Crystal Structure	Training Set			Test Set		
	RMSE/eV	MAE/eV	$R^2$	RMSE/eV	MAE/eV	$R^2$
FCC	0.03	0.02	0.99	0.15	0.11	0.89
BCC	0.10	0.07	0.98	0.22	0.17	0.85
HCP	0.09	0.07	0.97	0.23	0.18	0.76
FCC + BCC + HCP	0.11	0.08	0.97	0.20	0.16	0.91

Next, we studied the relative importance of features (out of total 20 features) used in all statistical models including the GBR model. Figure 2B showed the feature-importance scores of all 20 descriptors. We found that host metal's d-shell was the least important feature and its contribution towards the

prediction of surface segregation energy was negligible, while all other features contribute in model predictions. The 11 most significant features identified in this work were guest metal's work function, guest metal's electron affinity, host metal's density, host metal's atomic mass, guest metal's density, guest metal's atomic number, host metal's atomic number, guest metal's ionization energy, guest metal's radius, host metal's electron affinity and guest metal's Pauling electronegativity. Subsequently, we developed the GBR model using the 11 most significant features (Figure 3A) and obtained RMSE values of 0.04 eV and 0.14 eV for training and test sets (Table 3), respectively. These values were comparable to the values obtained using 20 features. We also investigated the predicting accuracy of the GBR model using the 5 most important features (guest metal's work function, host metal's atomic number, guest metal's electron affinity, host metal's atomic mass and guest metal's density; Figure 3B) and obtained an RMSE of 0.05 eV and 0.14 eV for training and test sets (Table 4), respectively. This value of RMSE was very similar to the values obtained using 20 and 11 features. Thus, we concluded that 5 features could be a reasonable choice for the prediction of surface segregation energies on (111) surface of bimetallic alloys using the GBR model.



**Figure 3.** GBR model predictions for FCC(111) surfaces using (A) 11 features and (B) 5 features. SSE = surface segregation energy.

**Table 3.** RMSE, MAE and  $R^2$  values for training and test sets obtained from the GBR model using 11 features.

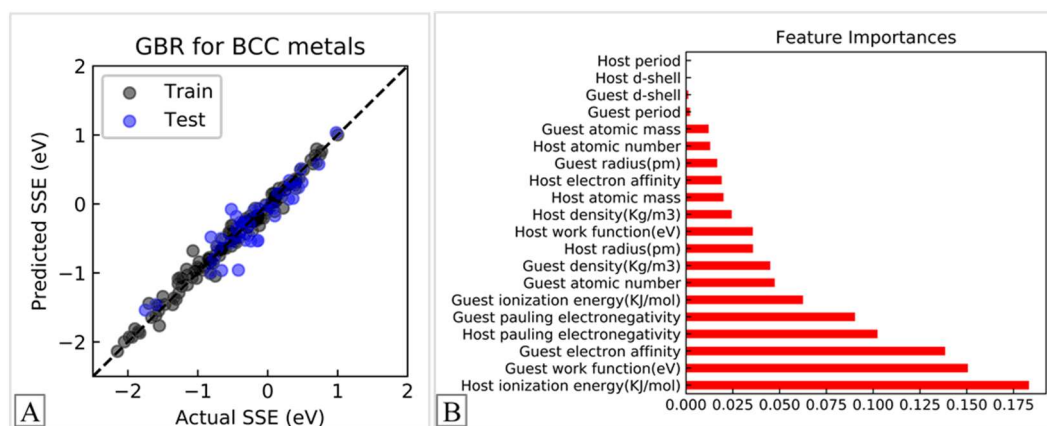
Crystal Structure	Training Set			Test Set		
	RMSE	MAE	$R^2$	RMSE	MAE	$R^2$
FCC	0.04	0.03	0.99	0.14	0.10	0.89
BCC	0.10	0.07	0.98	0.22	0.18	0.84
HCP	0.10	0.08	0.96	0.24	0.19	0.75
FCC + BCC + HCP	0.11	0.09	0.97	0.22	0.17	0.90

**Table 4.** RMSE, MAE and  $R^2$  values for training and test sets obtained from the GBR model using 5 features.

Crystal Structure	Training Set			Test Set		
	RMSE/eV	MAE/eV	$R^2$	RMSE/eV	MAE/eV	$R^2$
FCC	0.05	0.04	0.99	0.14	0.12	0.88
BCC	0.12	0.10	0.97	0.27	0.20	0.78
HCP	0.12	0.09	0.95	0.25	0.20	0.73
FCC + BCC + HCP	0.15	0.12	0.94	0.24	0.19	0.88

### 3.2. Model Predictions of the Segregation Energy on Bimetallic BCC (110) Surfaces

Machine learning models, OLS, PLS, SVR, GPR, GBR and KRR, similar to those used on FCC (111) surfaces, were used to model the surface segregation energies on bimetallic BCC (110) surfaces. The results demonstrated that the GBR outperformed other models (Figure 4A), having an RMSE of 0.22 eV for the test set and 0.10 eV for the training set. The estimated (MAE) of our training and test sets, were 0.07 eV and 0.17 eV, respectively. Moreover, the results presented in Table 2 show that the  $R^2$  values for the training and test sets were 0.98 and 0.85, respectively. Our current GBR model showed that the predicted surface segregation energy ( $-0.96$  eV) of FePt bimetallic system had the largest deviation from actual value ( $-0.66$  eV). Conversely, the RMSE and MAE values that were retrieved from other machine learning models (Figures S6–S10; Tables S1–S6) were higher as compared to those obtained using the GBR method. Also, the predictive ability of the GBR method and other methods considered was significantly better for the training set as compared to the test set. It was noteworthy that our GBR model essentially captured the sign of surface segregation energies on (110) surfaces of all bimetallic alloys. Therefore, the results suggested that the GBR model could be potentially used to predict surface segregation energies of BCC bimetallic (110) surfaces using the fundamental properties of transition metal components in alloys.



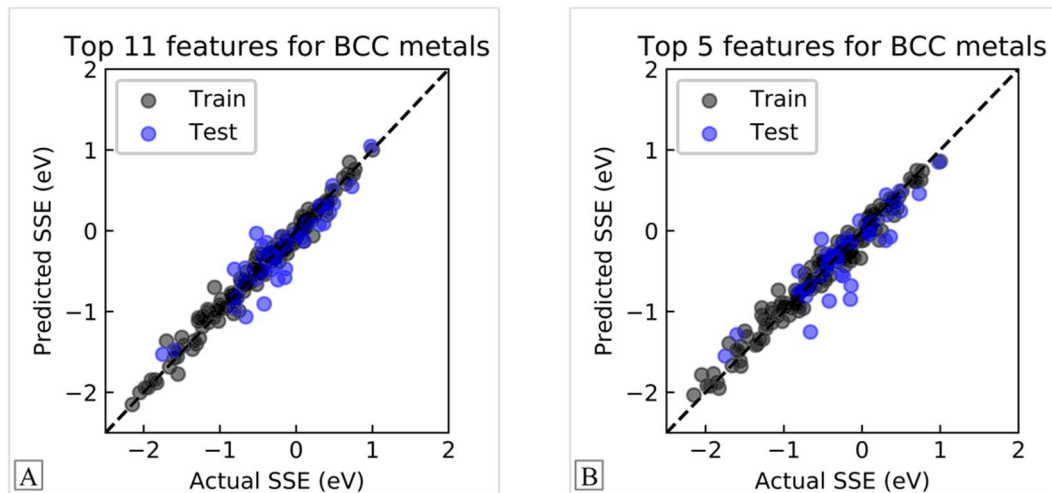
**Figure 4.** (A) Model predictions and (B) Feature importance scores for BCC(110) surfaces obtained from the GBR model. SSE = surface segregation energy.

The feature importance analysis on BCC (110) surfaces based on the GBR method was performed similarly compared to FCC (111) surfaces and the results are presented in Figure 4B. The results showed that the host d shell and host period were the least important features and their influence on the prediction of surface segregation energy was insignificant, while all other features contribute in model predictions. In line with the predictions on FCC (111) surfaces, we found that the model accuracy using the 11 topmost features (Figure 5A, Table 3: (RMSE; MAE) are (0.10 eV; 0.22 eV) and (0.07 eV; 0.18 eV) for the test and training sets, respectively) was similar to that obtained using all 20 features. In contrast, the model performance using the 5 topmost features (Figure 5B, Table 4: RMSE = 0.12 eV and 0.27 eV; MAE = 0.10 eV and 0.20 eV for the training and test sets, respectively) was slightly weaker compared to the models using 20 and 11 features.

### 3.3. Model Predictions of Segregation Energy for Bimetallic HCP (001) Surfaces

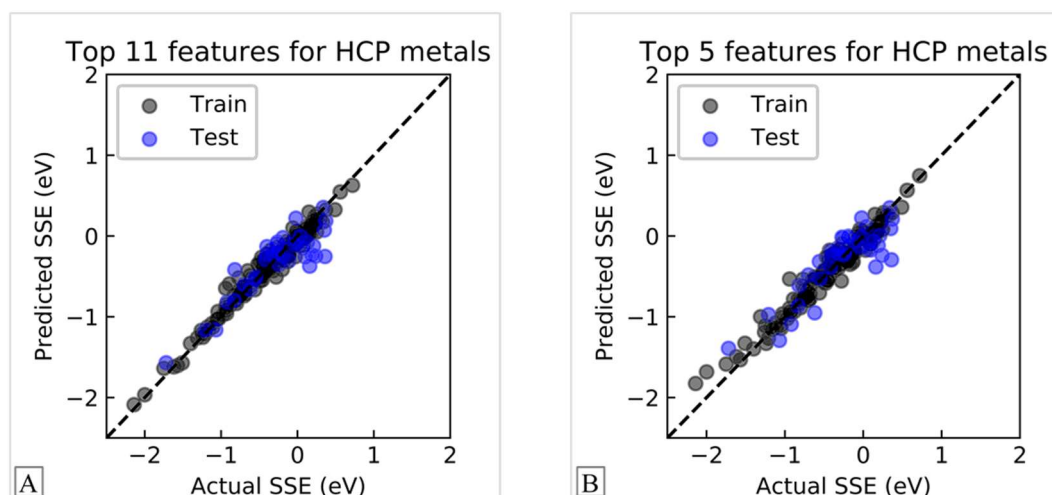
Similar linear and non-linear machine learning models as described in preceding sections were employed to predict surface segregation energies on bimetallic HCP (001) surfaces. Consistent with our results on FCC (111) and BCC (110) surfaces, the GBR model was superior and had (RMSE; MAE) values of (0.23 eV; 0.18 eV) for the test set and (0.09 eV; 0.07 eV) for the training set. Other models performed weakly and had higher RMSE and MAE values (Figures S11–S15; Tables S1–S6). Also, as presented in Table 2, the  $R^2$  values for the training and test sets were 0.97 and 0.76, respectively.

Once again, our GBR model efficiently captured the sign of surface segregation energies on (001) surfaces of all bimetallic alloys. RuCr and CoCr were found to be the outliers with largest deviation from the actual values. Our model was qualitatively able to capture the surface segregation behavior even though, the RMSE and MAE values of GBR model, the best performing model, were relatively higher compared to those estimated on FCC (111) and BCC (110) surfaces (Table 2).

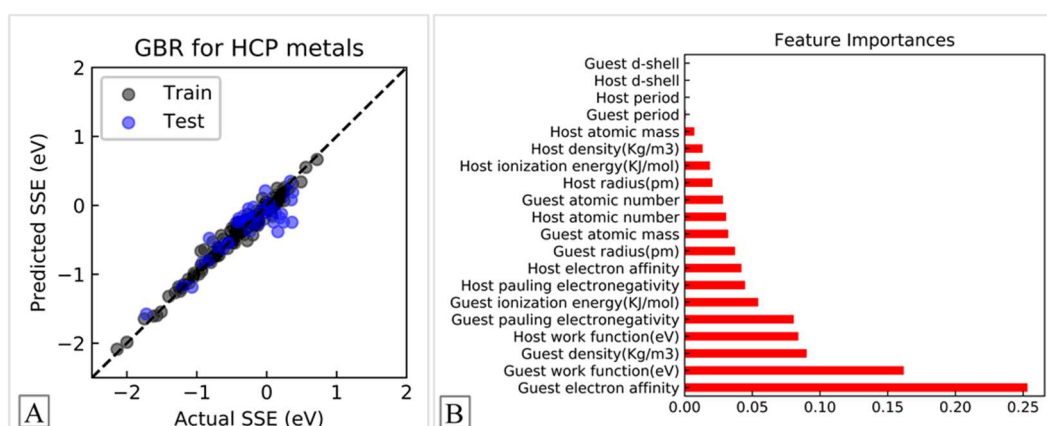


**Figure 5.** GBR model predictions for BCC(110) surfaces using (A) 11 features and (B) 5 features. SSE = surface segregation energy.

The GBR model using the most important 11 features (Figure 6A), out of 20 features (Figure 7B), has (RMSE; MAE) values of (0.10 eV; 0.08 eV) and (0.24 eV; 0.19 eV) for the training and test sets, respectively. It can be observed that the RMSE value for the test set (0.24 eV) from the GBR model with the 11 most significant features was slightly higher than the RMSE value (0.23 eV) that was obtained when all 20 features were used in the model. Additional analysis was performed to examine the predicting accuracy of our model by employing the 5 most important features (Figure 6B) and we obtained an (RMSE; MAE) value of (0.12 eV; 0.09 eV) for the training set and (0.25 eV; 0.20 eV) for the test set, respectively. The smaller difference of the values of RMSE and MAE between the models using various number of features suggested that the choice of 5 features could be a reasonable choice to qualitatively predict the surface segregation energies on HCP (001) surfaces.



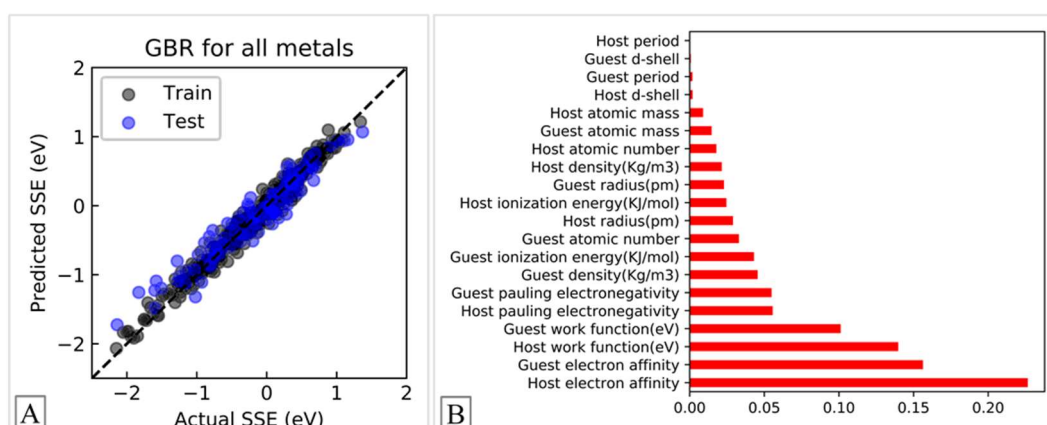
**Figure 6.** GBR model predictions for HCP(001) surfaces using (A) 11 features and (B) 5 features. SSE = surface segregation energy.



**Figure 7.** (A) Model predictions and (B) Feature importance scores for HCP(001) surfaces obtained from the GBR model. SSE = surface segregation energy.

### 3.4. Model Predictions of the Segregation Energy on Combining Bimetallic FCC (111), BCC (110) and HCP(001) Surfaces] Dataset

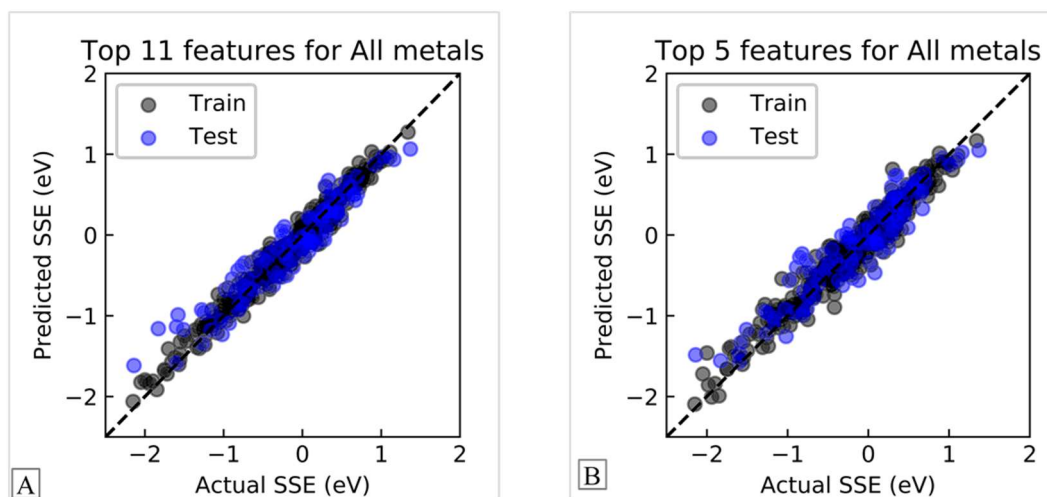
Finally, OLS, PLS, SVR, GPR, GBR and KRR models were developed for a combined data set (that contains surface segregation energies on FCC (111), BCC (110) and HCP (001) surfaces). It was found that the GBR performed better than other models (Figure 8A) with an RMSE of 0.11 eV for the training set and 0.20 eV for the test set. The MAE of our training and testing set were estimated to be 0.08 eV and 0.16 eV, respectively. As presented in Table 2, the  $R^2$  values for the training and testing sets were 0.97 and 0.91, respectively. We found that the RMSE and MAE values from other machine learning models (Figures S16–S20; Tables S1–S6) were higher as compared to those obtained from our GBR model. We found similar feature importance in the combined data set (Figure 8B) compared to individual FCC (111), BCC (110) and HCP (001) datasets. Thus, 5 topmost features for which the RMSE and MAE values were slightly higher to those obtained using 20 and 11 features and is a reasonable choice to predict qualitative trends of surface segregation energies on low index bimetallic surfaces (Figure 9A,B).



**Figure 8.** (A) Model predictions and (B) Feature importance scores for combined [FCC (111) + BCC (110) + HCP (001) surfaces] dataset obtained from the GBR model. SSE = surface segregation energy.

Overall, our calculations showed that the surface segregation energies on low index bimetallic surfaces can be correlated with elemental properties and predicted using simple statistical models. The predictive performance of simple non-linear statistical methods such as GBR was remarkable given the small size of the data set. The predictability of such models is expected to improve for a larger data set and thus, the future data-fitting strategy should focus on increasing the size of the input

dataset. Some of the features in input features are related to each other. Subsequently, we performed additional analysis removing (d-shell), (d-shell + atomic radius + volume) and (d-shell + atomic radius + volume) and adding a new feature  $\text{volume} = 4r^3\pi/3$  and the results are presented as Tables S7–S9 and Figures S26–S37 in the SI. The new results showed similar RMSE and MAE values for the GBR model. These observations suggested that some of the features can be combined to reduce the number of input features.



**Figure 9.** GBR model predictions for combined [FCC (111) + BCC (110) + HCP (001) surfaces] dataset using (A) 11 features and (B) 5 features. SSE = surface segregation energy.

#### 4. Conclusions

In summary, we used various linear and non-linear regression methods to train the surface segregation energies calculated using Green's-function linear-muffin-tin-orbitals method and correlate with the elemental properties of transition metals. The results showed that the surface segregation energy of the bimetallic alloys can be moderately predicted using simple machine learning models. It was noted that the non-linear gradient boosting regression (GBR) outperformed all other statistical methods employed in the present study. Using 20 features, the RMSE and MAE values for training/test sets in GBR model were 0.03/0.15 eV and 0.02/0.11 eV for FCC (111), 0.10/0.22 eV and 0.07/0.17 eV for BCC (110) and 0.09/0.23 eV and 0.07/0.18 eV for HCP (001) surfaces, respectively. Our feature importance analysis indicated that the 11 most significant features (common features include: guest work function, guest electron affinity, host density, host atomic mass, guest density, guest atomic number, host atomic number, guest ionization energy, guest radius, host electron affinity and guest Pauling electronegativity) produced similar results compared to those obtained by using all 20 features. The overall performance of the GBR model using the topmost 5 features is slightly weaker compared to the models using 20 and the topmost 11 features. This study, therefore, showed that the surface segregation energies on low index bimetallic surfaces can be predicted using elemental properties given the enough DFT database is available.

**Supplementary Materials:** The following are available online at <http://www.mdpi.com/1996-1073/13/9/2182/s1>, Figure S1–S37, Tables S1–S10.

**Author Contributions:** D.O. performed the statistical analysis. D.O. and S.K. co-wrote the manuscript. S.K. supervised the project. All authors have read and agreed to the published version of the manuscript.

**Acknowledgments:** S.K. acknowledges the faculty start-up fund from Florida A&M University.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Cheah, K.W.; Taylor, M.J.; Osatiashtiani, A.; Beaumont, S.K.; Nowakowski, D.J.; Yusup, S.; Bridgwater, A.V.; Kyriakou, G. Monometallic and Bimetallic Catalysts Based on Pd, Cu and Ni for Hydrogen Transfer Deoxygenation of a Prototypical Fatty Acid to Diesel Range Hydrocarbons. *Catal. Today* **2019**. [[CrossRef](#)]
2. Bizkarra, K.; Bermudez, J.M.; Arcelus-Arillaga, P.; Barrio, V.L.; Cambra, J.F.; Millan, M. Nickel Based Monometallic and Bimetallic Catalysts for Synthetic and Real Bio-Oil Steam Reforming. *Int. J. Hydrogen Energy* **2018**, *43*, 11706–11718. [[CrossRef](#)]
3. Greeley, J.; Stephens, I.E.L.; Bondarenko, A.S.; Johansson, T.P.; Hansen, H.A.; Jaramillo, T.F.; Rossmeisl, J.; Chorkendorff, I.; Nørskov, J.K. Alloys of Platinum and Early Transition Metals as Oxygen Reduction Electrocatalysts. *Nat. Chem.* **2009**, *1*, 552–556. [[CrossRef](#)] [[PubMed](#)]
4. Ford, D.C.; Nilekar, A.U.; Xu, Y.; Mavrikakis, M. Partial and Complete Reduction of O<sub>2</sub> by Hydrogen on Transition Metal Surfaces. *Surf. Sci.* **2010**, *604*, 1565–1575. [[CrossRef](#)]
5. Singh, A.; Spiccia, L. Water Oxidation Catalysts Based on Abundant 1st Row Transition Metals. *Coord. Chem. Rev.* **2013**, *257*, 2607–2622. [[CrossRef](#)]
6. Yu, W.; Porosoff, M.D.; Chen, J.G. Review of Pt-Based Bimetallic Catalysis: From Model Surfaces to Supported Catalysts. *Chem. Rev.* **2012**, *112*, 5780–5817. [[CrossRef](#)]
7. Gandeepan, P.; Müller, T.; Zell, D.; Cera, G.; Warratz, S.; Ackermann, L. 3d Transition Metals for C-H Activation. *Chem. Rev.* **2019**, *119*, 2192–2452. [[CrossRef](#)]
8. Duan, Z.; Wang, G. A first principles study of oxygen reduction reaction on a Pt (111) surface modified by a subsurface transition metal M (M= Ni, Co, or Fe). *Phys. Chem. Chem. Phys.* **2011**, *13*, 20178–20187. [[CrossRef](#)]
9. Han, X.; Zhou, R.; Lai, G.; Zheng, X. Influence of Support and Transition Metal (Cr, Mn, Fe, Co, Ni and Cu) on the Hydrogenation of p-Chloronitrobenzene over Supported Platinum Catalysts. *Catal. Today* **2004**, *93–95*, 433–437. [[CrossRef](#)]
10. She, Z.W.; Kibsgaard, J.; Dickens, C.F.; Chorkendorff, I.; Nørskov, J.K.; Jaramillo, T.F. Combining Theory and Experiment in Electrocatalysis: Insights into Materials Design. *Science* **2017**, *355*, eaad4998.
11. Stamenkovic, V.R.; Fowler, B.; Mun, B.S.; Wang, G.; Ross, P.N.; Lucas, C.A.; Markovic, N.M. Improved Oxygen Reduction Activity on Pt<sub>3</sub>Ni(111) via Increased Surface Site Availability. *Science* **2007**, *315*, 493–497. [[CrossRef](#)] [[PubMed](#)]
12. Greeley, J.; Mavrikakis, M. Alloy Catalysts Designed from First Principles. *Nat. Mater.* **2004**, *3*, 810–815. [[CrossRef](#)] [[PubMed](#)]
13. Vetere, V.; Merlo, A.B.; Ruggera, J.F.; Casella, M.L. Transition Metal-Based Bimetallic Catalysts for the Chemoselective Hydrogenation of Furfuraldehyde. *J. Braz. Chem. Soc.* **2010**, *21*, 914–920. [[CrossRef](#)]
14. Song, C.; Zhang, J. Electrocatalytic Oxygen Reduction Reaction. In *PEM Fuel Cell Electrocatalysts and Catalyst Layers: Fundamentals and Applications*; Springer: London, UK, 2008; pp. 89–134.
15. Stamenkovic, V.R.; Mun, B.S.; Arenz, M.; Mayrhofer, K.J.J.; Lucas, C.A.; Wang, G.; Ross, P.N.; Markovic, N.M. Trends in Electrocatalysis on Extended and Nanoscale Pt-Bimetallic Alloy Surfaces. *Nat. Mater.* **2007**, *6*, 241–247. [[CrossRef](#)] [[PubMed](#)]
16. Long, J.; Shen, K.; Chen, L.; Li, Y. Multimetal-MOF-Derived Transition Metal Alloy NPs Embedded in an N-Doped Carbon Matrix: Highly Active Catalysts for Hydrogenation Reactions. *J. Mater. Chem. A* **2016**, *4*, 10254–10262. [[CrossRef](#)]
17. Fernandes, R.; Patel, N.; Miotello, A.; Jaiswal, R.; Kothari, D.C. Dehydrogenation of Ammonia Borane with Transition Metal-Doped Co-B Alloy Catalysts. *Int. J. Hydrog. Energy* **2012**, *37*, 2397–2406. [[CrossRef](#)]
18. Sankar, M.; Dimitratos, N.; Miedziak, P.J.; Wells, P.P.; Kiely, C.J.; Hutchings, G.J. Designing Bimetallic Catalysts for a Green and Sustainable Future. *Chem. Soc. Rev.* **2012**, *41*, 8099–8139. [[CrossRef](#)]
19. Liao, H.; Fisher, A.; Xu, Z.J. Surface Segregation in Bimetallic Nanoparticles: A Critical Issue in Electrocatalyst Engineering. *Small* **2015**, *11*, 3221–3246. [[CrossRef](#)]
20. Sachtler, W.M.H. Surface Composition of Alloys. *Appl. Surf. Sci.* **1984**, *19*, 167–180. [[CrossRef](#)]
21. Prasada Rao, T.S.R. Surface Enrichment in Catalysts. *Catal. Rev.* **1979**, *20*, 97–120.
22. Wolf, R.M.; Siera, J.; Van Delft, F.C.M.J.M.; Nieuwenhuys, B.E. A Comparative Study of the Behaviour of Single-Crystal Surfaces and Supported Catalysts in NO Reduction and CO Oxidation over Pt-Rh Alloys. *Faraday Discuss. Chem. Soc.* **1989**, *87*, 275–289. [[CrossRef](#)]

23. Sachtler, W.M.H.; van Santen, R.A. Surface Composition of Binary Alloys. *Appl. Surf. Sci.* **1979**, *3*, 121–144. [[CrossRef](#)]
24. Wang, G.; Van Hove, M.A.; Ross, P.N.; Baskes, M.I. Monte Carlo Simulations of Segregation in Pt-Ni Catalyst Nanoparticles. *J. Chem. Phys.* **2005**, *122*, 024706. [[CrossRef](#)] [[PubMed](#)]
25. Wang, G.; Van Hove, M.A.; Ross, P.N.; Baskes, M.I. Quantitative Prediction of Surface Segregation in Bimetallic Pt-M Alloy Nanoparticles (M = Ni, Re, Mo). *Prog. Surf. Sci.* **2005**, *79*, 28–45. [[CrossRef](#)]
26. Corcoran, C.J.; Tavassol, H.; Rigsby, M.A.; Bagus, P.S.; Wieckowski, A. Application of XPS to Study Electrocatalysts for Fuel Cells. *J. Power Sour.* **2010**, *195*, 7856–7879. [[CrossRef](#)]
27. Olovsson, W.; Holmström, E.; Sandell, A.; Abrikosov, A. Core-Level Shifts for Surface Bimetallic Systems from First-Principles Theory: Pd-Mn Structures on Pd(100). *Phys. Rev. B Condens. Matter Mater. Phys.* **2003**, *68*, 045411. [[CrossRef](#)]
28. Medford, A.J.; Vojvodic, A.; Hummelshøj, J.S.; Voss, J.; Abild-Pedersen, F.; Studt, F.; Bligaard, T.; Nilsson, A.; Nørskov, J.K. From the Sabatier Principle to a Predictive Theory of Transition-Metal Heterogeneous Catalysis. *J. Catal.* **2015**, *328*, 36–42. [[CrossRef](#)]
29. Yu, Y.; Xiao, W.; Wang, J.; Wang, L. Understanding the Surface Segregation Behavior of Transition Metals on Ni(111): A First-Principles Study. *Phys. Chem. Chem. Phys.* **2016**, *18*, 26616–26622. [[CrossRef](#)]
30. Farsi, L.; Deskins, N.A. First Principles Analysis of Surface Dependent Segregation in Bimetallic Alloys. *Phys. Chem. Chem. Phys.* **2019**, *21*, 23626–23637. [[CrossRef](#)]
31. Boom, R.; de Boer, F.R. *Energy Effects in Bulk Metals. Encyclopedia of Materials: Science and Technology*; Elsevier: Pergamon, Greece, 2006; pp. 1–7.
32. Ruban, A.V.; Skriver, H.L.; Nørskov, J.K. Surface Segregation Energies in Transition-Metal Alloys. *Phys. Rev. B Condens. Matter Mater. Phys.* **1999**, *59*, 15990–16000. [[CrossRef](#)]
33. Nilekar, A.U.; Ruban, A.V.; Mavrikakis, M. Surface Segregation Energies in Low-Index Open Surfaces of Bimetallic Transition Metal Alloys. *Surf. Sci.* **2009**, *603*, 91–96. [[CrossRef](#)]
34. Ruban, A.V.; Skriver, H.L. Calculated Surface Segregation in Transition Metal Alloys. *Comput. Mater. Sci.* **1999**, *15*, 119–143. [[CrossRef](#)]
35. Ma, Y.; Balbuena, P.B. Pt Surface Segregation in Bimetallic Pt3M Alloys: A Density Functional Theory Study. *Surf. Sci.* **2008**, *602*, 107–113. [[CrossRef](#)]
36. Zhang, Y.; Duan, Z.; Xiao, C.; Wang, G. Density Functional Theory Calculation of Platinum Surface Segregation Energy in Pt3Ni (111) Surface Doped with a Third Transition Metal. *Surf. Sci.* **2011**, *605*, 1577–1582. [[CrossRef](#)]
37. Zahrt, A.F.; Henle, J.J.; Rose, B.T.; Wang, Y.; Darrow, W.T.; Denmark, S.E. Prediction of Higher-Selectivity Catalysts by Computer-Driven Workflow and Machine Learning. *Science* **2019**, *363*, eaau5631. [[CrossRef](#)]
38. Li, Z.; Wang, S.; Chin, W.S.; Achenie, L.E.; Xin, H. High-Throughput Screening of Bimetallic Catalysts Enabled by Machine Learning. *J. Mater. Chem. A* **2017**, *5*, 24131–24138. [[CrossRef](#)]
39. Schlexer Lamoureux, P.; Winther, K.T.; Garrido Torres, J.A.; Streibel, V.; Zhao, M.; Bajdich, M.; Abild-Pedersen, F.; Bligaard, T. Machine Learning for Computational Heterogeneous Catalysis. *ChemCatChem* **2019**, *11*, 3581–3601. [[CrossRef](#)]
40. Yang, W.; Fidelis, T.T.; Sun, W.H. Machine Learning in Catalysis, from Proposal to Practicing. *ACS Omega* **2020**, *5*, 83–88. [[CrossRef](#)]
41. Toyao, T.; Maeno, Z.; Takakusagi, S.; Kamachi, T.; Takigawa, I.; Shimizu, K.I. Machine Learning for Catalysis Informatics: Recent Applications and Prospects. *ACS Catal.* **2020**, *10*, 2260–2297. [[CrossRef](#)]
42. Ma, X.; Li, Z.; Achenie, L.E.K.; Xin, H. Machine-Learning-Augmented Chemisorption Model for CO<sub>2</sub> Electroreduction Catalyst Screening. *J. Phys. Chem. Lett.* **2015**, *6*, 3528–3533. [[CrossRef](#)]
43. Ulissi, Z.W.; Medford, A.J.; Bligaard, T.; Nørskov, J.K. To Address Surface Reaction Network Complexity Using Scaling Relations Machine Learning and DFT Calculations. *Nat. Commun.* **2017**, *8*, 14621. [[CrossRef](#)]
44. Tran, K.; Ulissi, Z.W. Active Learning across Intermetallics to Guide Discovery of Electrocatalysts for CO<sub>2</sub> Reduction and H<sub>2</sub> Evolution. *Nat. Catal.* **2018**, *1*, 696–703. [[CrossRef](#)]
45. Noh, J.; Back, S.; Kim, J.; Jung, Y. Active Learning with Non-: Ab Initio Input Features toward Efficient CO<sub>2</sub> Reduction Catalysts. *Chem. Sci.* **2018**, *9*, 5152–5159. [[CrossRef](#)]
46. HaŁas, S. 100 Years of Work Function. *Mater. Sci.-Poland* **2006**, *24*, 951–966.
47. William, D.; Callister, D.G.R., Jr. *Materials Science and Engineering: An Introduction*, 8th ed.; John Wiley and Sons: Hoboken, NJ, USA, 1994.

48. Wold, S.; Sjöström, M.; Eriksson, L. PLS-Regression: A Basic Tool of Chemometrics. *Chemom. Intell. Lab. Syst.* **2001**, *58*, 109–130. [[CrossRef](#)]
49. Friedman, J.H. Stochastic Gradient Boosting. *Comput. Stat. Data Anal.* **2002**, *38*, 367–378. [[CrossRef](#)]
50. Schulz, E.; Speekenbrink, M.; Krause, A. A Tutorial on Gaussian Process Regression: Modelling, Exploring, and Exploiting Functions. *J. Math. Psychol.* **2018**, *85*, 1–16. [[CrossRef](#)]
51. Vovk, V. *Kernel Ridge Regression. Empirical Inference: Festschrift in Honor of Vladimir N. Vapnik*; Springer: Heidelberg Germany, 2013; pp. 105–116.
52. Smola, A.J.; Schölkopf, B. A Tutorial on Support Vector Regression. *Stat. Comput.* **2004**, *14*, 199–222. [[CrossRef](#)]
53. Scikit-Learn: Machine Learning in Python. Available online: <https://scikit-learn.org> (accessed on 27 April 2020).



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).