

## Interpretable Machine-Learning Prediction of DFT Energies per Atom and Identification of Magic Numbers in Coinage-Metal Nanoclusters ( $N \leq 55$ ) from the Open Quantum Cluster Database

Ali A Khairbek<sup>1\*</sup>, Maha I. Al-Zaben<sup>2</sup>, Abdullah Yahya Abdullah Alzahrani<sup>3</sup>, and Renjith Thomas<sup>4,5\*</sup>

<sup>1</sup> Department of Mechanical Engineering, University Centre for Research & Development, Chandigarh University, Gharuan, Mohali, Punjab, India

<sup>2</sup> Department of Chemistry, College of Science, King Saud University, P.O. Box 2455, Riyadh 11451, Saudi Arabia

<sup>3</sup> Department of Chemistry, Faculty of Science, King Khalid University, Abha, 61413, Saudi Arabia

<sup>4</sup> Department of Chemistry, St Berchmans College (Autonomous), Mahatma Gandhi University, Changanassery, Kerala, India-686101.

<sup>5</sup> Centre for Theoretical and Computational Chemistry, St Berchmans College (Autonomous), Changanassery, Kerala, India-686101.

Corresponding Author: **Prof. Renjith Thomas** [renjith@sbcollege.ac.in](mailto:renjith@sbcollege.ac.in), **Dr. Ali A Khairbek** [alikhairbek@gmail.com](mailto:alikhairbek@gmail.com)

### Contents

**S1.** Extended ML vs DFT comparison for all cluster sizes (Table S1)

**S2.** Per-model cross-validation results (Table S2)

**S3.** Per-metal model accuracy (Table S3)

**S4.** Full-feature SHAP importance ranking (Table S4)

**S5.** Geometry-only model: full benchmarking (Table S5)

**S6.** Per-metal magic-number tables with full  $\Delta^2E$  series (Tables S6–S8)

**S7.** Reproducibility checklist and code availability

### S1. Extended ML vs DFT Comparison for All Cluster Sizes

**Table S1.** Average DFT and ML-predicted energy per atom (eV/atom) and per-size mean absolute error (MAE) for all nuclearities  $N = 5-55$  in the test set. Values are recomputed with

the corrected pipeline (LightGBM, full feature set including cluster size N). The full numerical table is also provided as a supplementary CSV file.

<b>N</b>	<b>n samples</b>	<b>Avg DFT (eV/atom)</b>	<b>Avg ML (eV/atom)</b>	<b>MAE (eV/atom)</b>
5	7	-1.5679	-1.5999	0.0489
6	5	-1.7123	-1.7877	0.0754
7	11	-1.9573	-1.9406	0.0289
8	7	-1.9774	-1.9812	0.0172
9	14	-2.1244	-2.1103	0.0364
10	16	-2.0869	-2.0837	0.0275
11	8	-2.1498	-2.1519	0.0190
12	4	-2.2664	-2.2538	0.0168
13	16	-2.1895	-2.1854	0.0201
14	14	-2.2028	-2.2081	0.0198
15	13	-2.3614	-2.3478	0.0250
16	14	-2.1890	-2.1940	0.0168
17	14	-2.2670	-2.2709	0.0159
18	8	-2.3789	-2.3792	0.0173
19	10	-2.3384	-2.3371	0.0085
20	16	-2.4695	-2.4603	0.0182
21	13	-2.4316	-2.4300	0.0131
22	14	-2.3853	-2.3888	0.0176
23	19	-2.3786	-2.3782	0.0131
24	13	-2.5367	-2.5440	0.0218
25	20	-2.4844	-2.4849	0.0108
26	18	-2.4841	-2.4975	0.0204
27	6	-2.3824	-2.3873	0.0137
28	19	-2.4329	-2.4386	0.0118
29	6	-2.3474	-2.3452	0.0115

30	17	-2.5592	-2.5585	0.0074
31	17	-2.4458	-2.4433	0.0142
32	17	-2.4539	-2.4561	0.0083
33	14	-2.5465	-2.5446	0.0107
34	18	-2.4101	-2.4079	0.0086
35	12	-2.4563	-2.4669	0.0138
36	12	-2.4925	-2.4920	0.0087
37	15	-2.4700	-2.4752	0.0092
38	10	-2.5676	-2.5735	0.0087
39	7	-2.6921	-2.6945	0.0115
40	15	-2.5240	-2.5249	0.0065
41	13	-2.6661	-2.6660	0.0096
42	17	-2.4707	-2.4695	0.0059
43	18	-2.6220	-2.6197	0.0075
44	16	-2.6163	-2.6183	0.0056
45	12	-2.4586	-2.4579	0.0103
46	20	-2.6616	-2.6666	0.0119
47	6	-2.2569	-2.2558	0.0088
48	11	-2.6510	-2.6486	0.0085
49	13	-2.4115	-2.4122	0.0095
50	14	-2.6426	-2.6441	0.0120
51	9	-2.5750	-2.5814	0.0084
52	17	-2.5711	-2.5705	0.0080
53	7	-2.4509	-2.4555	0.0124
54	14	-2.4684	-2.4730	0.0109
55	12	-2.5444	-2.5489	0.0116

**Notes:** (i) values are derived from the test partition of the LightGBM model trained on the full feature set; (ii) per-size sample counts vary because the 70/15/15 split is stratified by

metal, not by size; (iii) per-size MAE remains uniformly below 0.08 eV/atom across the entire range, with the largest residuals at the smallest nuclearities ( $N = 5-7$ ) where structural variability is largest.

## S2. Per-Model Cross-Validation Results

**Table S2.** Detailed cross-validation performance for all seven regressors trained on the full feature set. Standard 5-fold KFold cross-validation provides a within-domain stability estimate; size-grouped CV (4 size bins:  $N = 3-15$ ,  $16-30$ ,  $31-45$ ,  $46-55$ ; one bin held out per fold) tests near-extrapolation across nuclearity domains.

Model	Test MAE	Test R <sup>2</sup>	5-fold CV MAE (mean)	5-fold CV MAE (std)	Size-Grouped CV MAE
LightGBM	0.01436	0.9957	0.01422	0.00044	0.107 ± 0.074
ExtraTrees	0.01453	0.9948	0.01389	0.00027	0.079 ± 0.081
CatBoost	0.01534	0.9947	0.01463	0.00039	0.079 ± 0.076
RandomForest	0.01548	0.9940	0.01507	0.00030	0.092 ± 0.072
GradientBoosting	0.01550	0.9943	0.01400	0.00061	0.086 ± 0.071
XGBoost	0.01597	0.9943	0.01457	0.00027	0.142 ± 0.071
NeuralNet (MLP)	0.02908	0.9748	0.02550	0.00112	0.078 ± 0.069

**Notes:** (i) all CV values are in eV/atom; (ii) the 30× smaller fold-to-fold standard deviations relative to the model-to-model spread indicate that the test-set MAE is a faithful estimator of generalisation accuracy within the QCD chemical and size domain; (iii) size-grouped CV degrades substantially across all models, quantifying the cost of out-of-domain extrapolation.

### S3. Per-Metal Model Accuracy

**Table S3.** Test-set absolute and squared error metrics broken down by metal, using the LightGBM full-feature model. Metal-stratified errors are useful for assessing whether the model treats Cu, Ag, and Au with comparable fidelity.

Metal	n samples	MAE (eV/atom)	RMSE (eV/atom)
Ag	232	0.01167	0.01706
Cu	192	0.01498	0.02463
Au	234	0.01652	0.02770

The slightly larger Au error is consistent with the broader range of low-energy Au isomers and the stronger sensitivity of relativistic systems to fine geometric variation.

### S4. Full-Feature SHAP Importance Ranking

**Table S4.** Mean absolute SHAP value for each input feature of the LightGBM full-feature model, sorted in descending order. SHAP values were computed on a representative sample of the test set; columns reflect the average magnitude of each feature's contribution to the predicted DFT energy per atom (eV/atom).

Rank	Feature	mean  SHAP  (eV/atom)	Category
1	metal_Z (atomic number)	0.27404	Composition
2	n_atoms (cluster size N)	0.15566	Size
3	compactness (Rg/d_mean)	0.00685	Geometric
4	homo_lumo_gap	0.00389	Electronic (DFT)
5	mean_dist (centroid distance)	0.00311	Geometric
6	std_dist	0.00277	Geometric
7	max_dist	0.00262	Geometric
8	bbox_z	0.00252	Geometric
9	asphericity	0.00230	Geometric
10	radius_gyration	0.00216	Geometric
11	magnetic_moment	0.00207	Electronic (DFT)

12	bbox_y	0.00148	Geometric
13	bbox_x	0.00108	Geometric
14	n_val_electrons	~0.0	Composition (redundant with N for monometallic clusters)

Two features dominate by two orders of magnitude over the others: metal identity ( $Z$ ) and cluster size ( $N$ ). Geometric and electronic descriptors enter at the  $\sim 10$  meV/atom scale and govern the small energy differences between isomers of the same size and metal that determine magic-number locations. Note: in this dataset  $n\_val\_electrons$  is exactly proportional to  $N$  for monoatomic Cu/Ag/Au clusters and therefore contributes negligibly when  $N$  is included.

### S5. Geometry-Only Model: Full Benchmarking

**Table S5.** Performance of all seven regressors trained on the geometry-only feature set (metal  $Z$ , cluster size  $N$ , and nine geometric descriptors, without DFT-derived electronic features). Same 70/15/15 stratified split as for the full-feature benchmark.

Model	MAE <sub>tes</sub> t	R <sup>2</sup> <sub>tes</sub> t	RMSE <sub>tes</sub> t	5-fold CV MAE	Size- Grouped CV MAE	Train (s)
LightGBM	0.01477	0.9955	0.0240	0.01419 ± 0.00069	0.110 ± 0.080	0.7
ExtraTrees	0.01513	0.9943	0.0271	0.01414 ± 0.00040	0.082 ± 0.081	2.4
RandomForest	0.01585	0.9942	0.0274	0.01529 ± 0.00033	0.094 ± 0.073	10.2
GradientBoosting	0.01649	0.9928	0.0305	0.01477 ± 0.00057	0.094 ± 0.075	20.8
CatBoost	0.01652	0.9938	0.0283	0.01539 ± 0.00038	0.087 ± 0.078	2.9
XGBoost	0.01871	0.9909	0.0344	0.01657 ± 0.00050	0.157 ± 0.073	1.5
NeuralNet (MLP)	0.03209	0.9694	0.0628	0.02851 ± 0.00090	0.082 ± 0.072	1.9

All metrics in eV/atom unless stated. Removing the two electronic features (HOMO–LUMO gap, magnetic moment) costs only  $\approx 0.4$ – $3.0$  meV/atom in MAE for the top three regressors, demonstrating that energies can be ranked accurately from coordinates alone within the QCD chemical and size domain.

## S6. Per-Metal Magic-Number Tables — Full $\Delta^2E$ Series

**Tables S6–S8.** Complete second-difference series  $\Delta^2E(N) = E(N+1) + E(N-1) - 2 E(N)$  for each coinage metal (Cu, Ag, Au), where  $E(N)$  is the average DFT energy per atom at size  $N$  within that metal. Adaptive thresholds: 0.0146, 0.0113, and 0.0112 eV/atom for Cu, Ag, and Au, respectively ( $= 0.5 \times \sigma(\Delta^2E)$ ). Sizes with  $\Delta^2E$  exceeding the threshold are highlighted in the “Detected” column. Literature values are taken from Häkkinen, de Heer, Knight, and Pyykkö (see main text).

**Table S6.** *Copper (Cu) —  $\Delta^2E$  series and detected magic numbers*

N	Avg E_DFT/N (eV/atom)	$\Delta^2E$ (eV/atom)	Detected?	Notes
5	-1.5934	—	—	
6	-1.9764	-0.058	No	Literature magic; not detected here
7	-2.1218	+0.005	No	
8	-2.2515	+0.048	Yes	Both (jellium $1S^21P^6$ )
10	-2.3621	+0.034	Yes	Detected only
12	-2.5074	+0.030	Yes	Detected only
14	-2.6238	+0.025	Yes	Detected only
18	-2.6843	+0.001	No	Literature magic; sub-threshold
20	-2.7308	+0.018	Yes	Both
24	-2.8094	+0.022	Yes	Detected only
34	-2.8971	+0.034	Yes	Both
37	-2.9189	+0.020	Yes	Detected only
40	-2.9437	+0.003	No	Literature magic; sub-threshold
55	-3.0124	—	—	

**Table S7.** Silver (Ag) —  $\Delta^2E$  series and detected magic numbers

<b>N</b>	<b>Avg E_DFT/N (eV/atom)</b>	<b><math>\Delta^2E</math> (eV/atom)</b>	<b>Detected ?</b>	<b>Notes</b>
5	-1.2890	—	—	
6	-1.4670	+0.039	Yes	Both (only metal where N=6 detected)
8	-1.6873	+0.110	Yes	Both (strongest peak)
10	-1.7964	+0.038	Yes	Detected only
12	-1.8612	+0.024	Yes	Detected only
18	-1.9607	+0.006	No	Literature magic; sub-threshold
20	-2.0032	+0.030	Yes	Both
24	-2.0588	+0.018	Yes	Detected only
26	-2.0712	+0.015	Yes	Detected only
32	-2.0987	+0.014	Yes	Detected only
34	-2.1136	+0.030	Yes	Both
37	-2.1284	+0.014	Yes	Detected only
40	-2.1444	+0.001	No	Literature magic; sub-threshold
53	-2.1887	+0.018	Yes	Detected only
55	-2.1968	—	—	Literature magic

**Table S8.** Gold (Au) —  $\Delta^2E$  series and detected magic numbers

<b>N</b>	<b>Avg E_DFT/N (eV/atom)</b>	<b><math>\Delta^2E</math> (eV/atom)</b>	<b>Detected ?</b>	<b>Notes</b>
5	-1.6789	—	—	
6	-1.8796	+0.026	Yes	Both
7	-1.9810	+0.014	Yes	Detected only
8	-2.0956	+0.091	Yes	Both (strongest peak)
10	-2.2090	+0.018	Yes	Detected only
12	-2.2987	+0.034	Yes	Detected only

14	-2.3431	+0.018	Yes	Detected only
16	-2.3830	+0.025	Yes	Detected only
18	-2.4165	+0.006	No	Literature magic; sub-threshold
19	-2.4286	+0.013	Yes	Detected only
20	-2.4465	-0.002	No	Literature magic; ANOMALOUS — see main text
25	-2.5078	+0.022	Yes	Detected only
28	-2.5448	+0.015	Yes	Detected only
32	-2.5844	+0.018	Yes	Both (Au-specific; relativistic)
34	-2.5942	+0.017	Yes	Both
36	-2.6082	+0.018	Yes	Detected only
38	-2.6327	+0.016	Yes	Both (Au-specific; relativistic)
53	-2.6956	+0.034	Yes	Detected only
55	-2.7023	—	—	Literature magic

## S8. Reproducibility Checklist

The following measures support full reproducibility of all results reported in the revised manuscript:

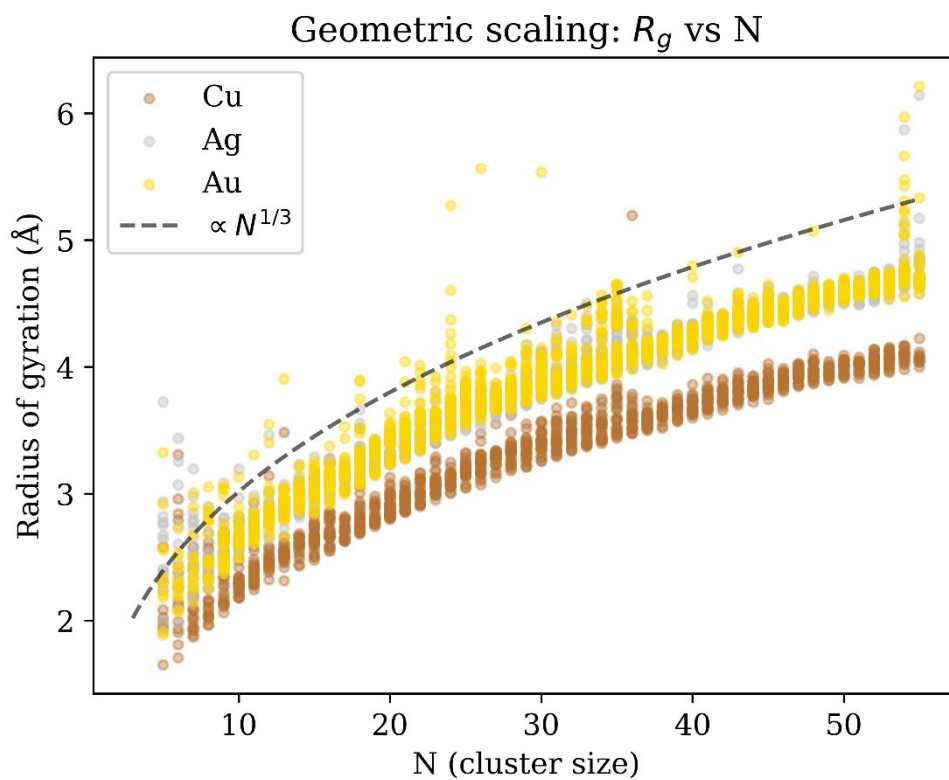
1. Random seeds fixed at 42 for all NumPy, scikit-learn, XGBoost, LightGBM, and CatBoost components.
2. Stratified 70/15/15 train/validation/test split implemented in `three_way_split()`, preserving the per-metal proportions in each subset.
3. 5-fold KFold cross-validation (`shuffle=True`, `random_state=42`) for all seven models on the full and geometry-only feature sets.
4. Size-grouped cross-validation with four nuclearity bins (3–15, 16–30, 31–45, 46–55).
5. Standardisation (zero mean, unit variance) fitted on training data only and applied to validation/test.
6. Adaptive metal-specific magic-number threshold:  $\Delta^2E_{\min} = 0.5 \times \sigma(\Delta^2E_{\text{metal}})$ .
7. All model hyperparameters listed explicitly in the open-source script.

8. Complete pipeline (Python + Jupyter notebook), trained models, processed dataset, all figures, and all tables openly available at <https://github.com/alikhairbek/CoinageClusterML>.

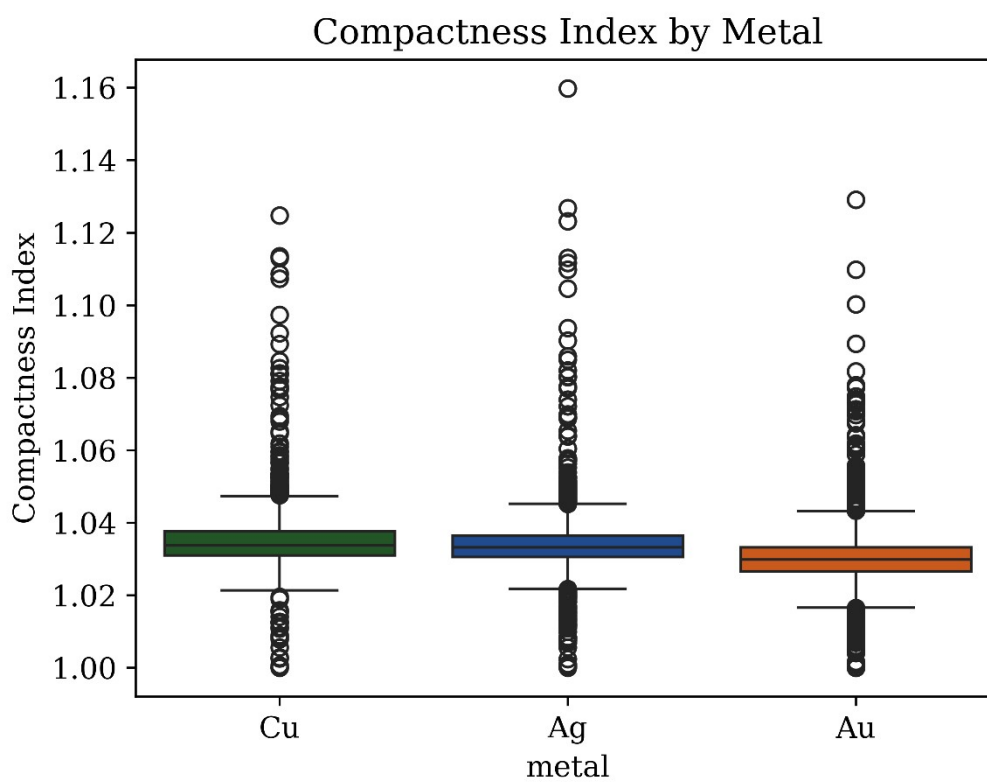
9. Code release date: revision submission to PCCP, R1.

### Software Versions

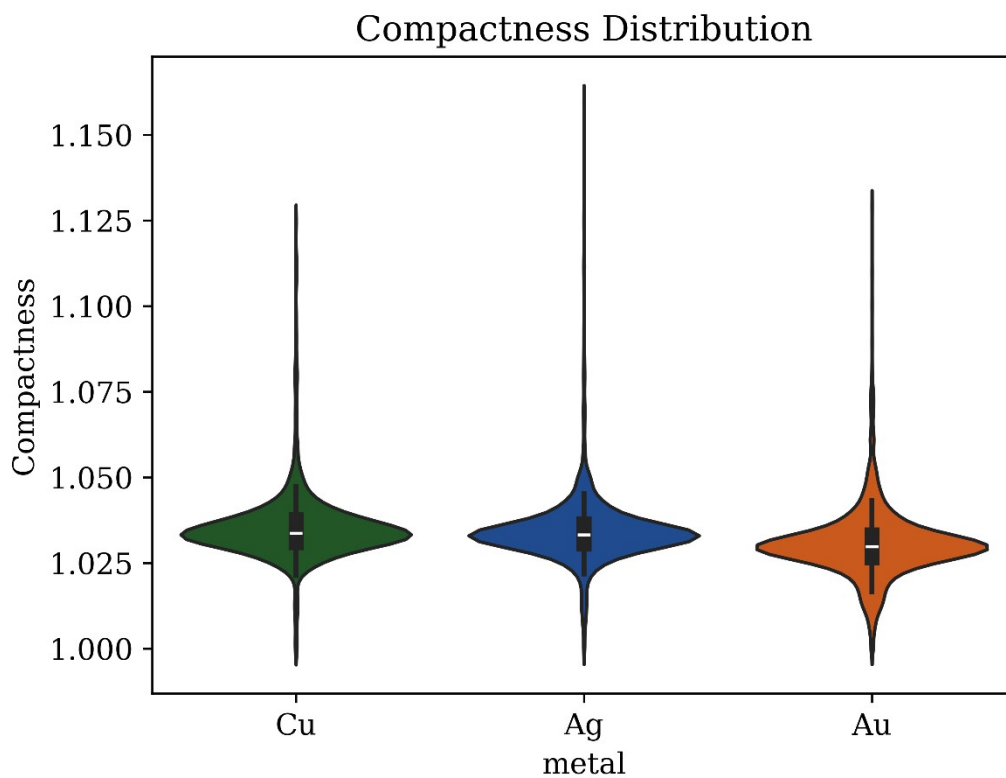
<b>Library</b>	<b>Version</b>
Python	3.12
NumPy	1.26
pandas	2.2
scikit-learn	1.4
XGBoost	2.0
LightGBM	4.3
CatBoost	1.2
SHAP	0.45
matplotlib	3.8
seaborn	0.13



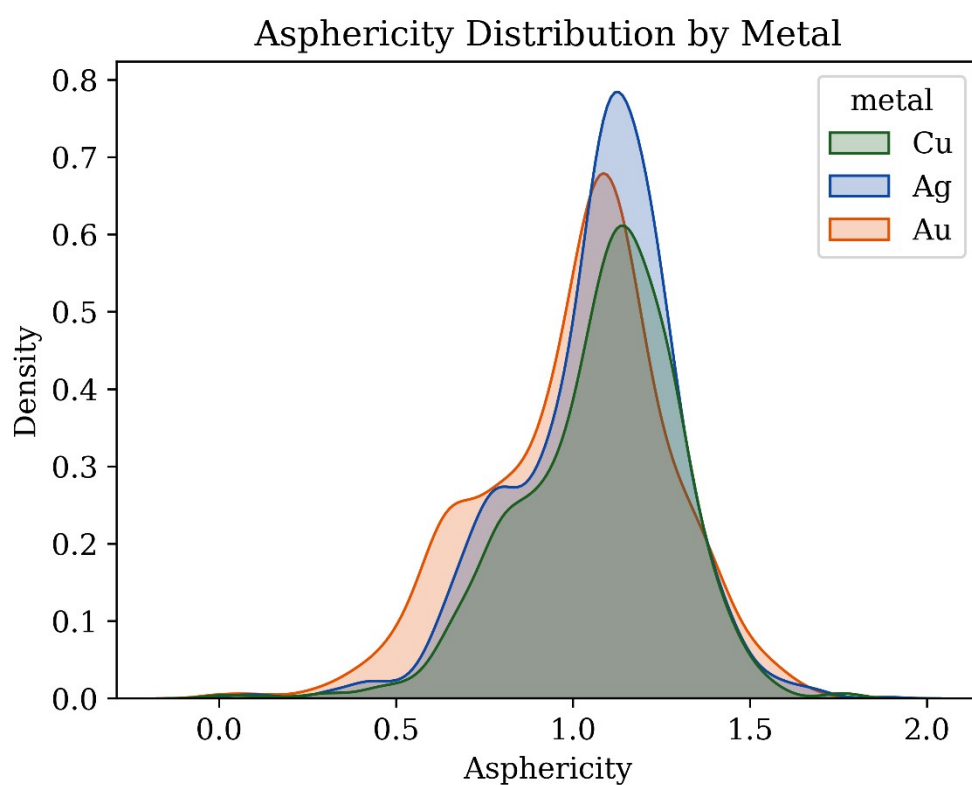
**Figure S1.** Radius of gyration vs cluster size (per metal)



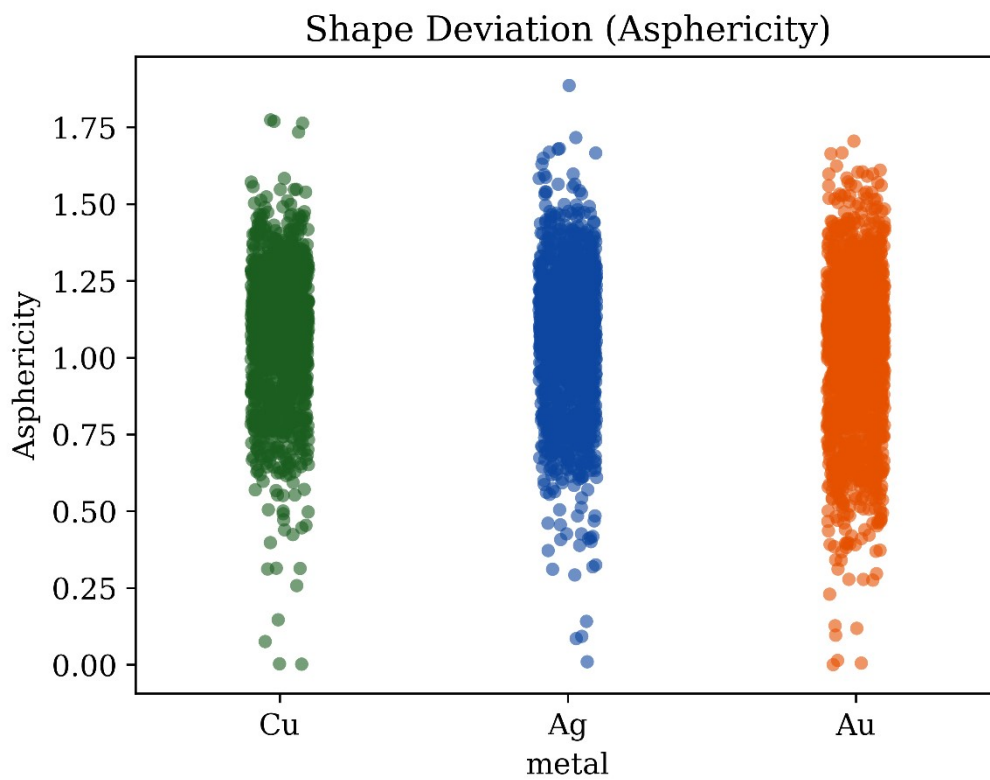
**Figure S2.** Compactness Index



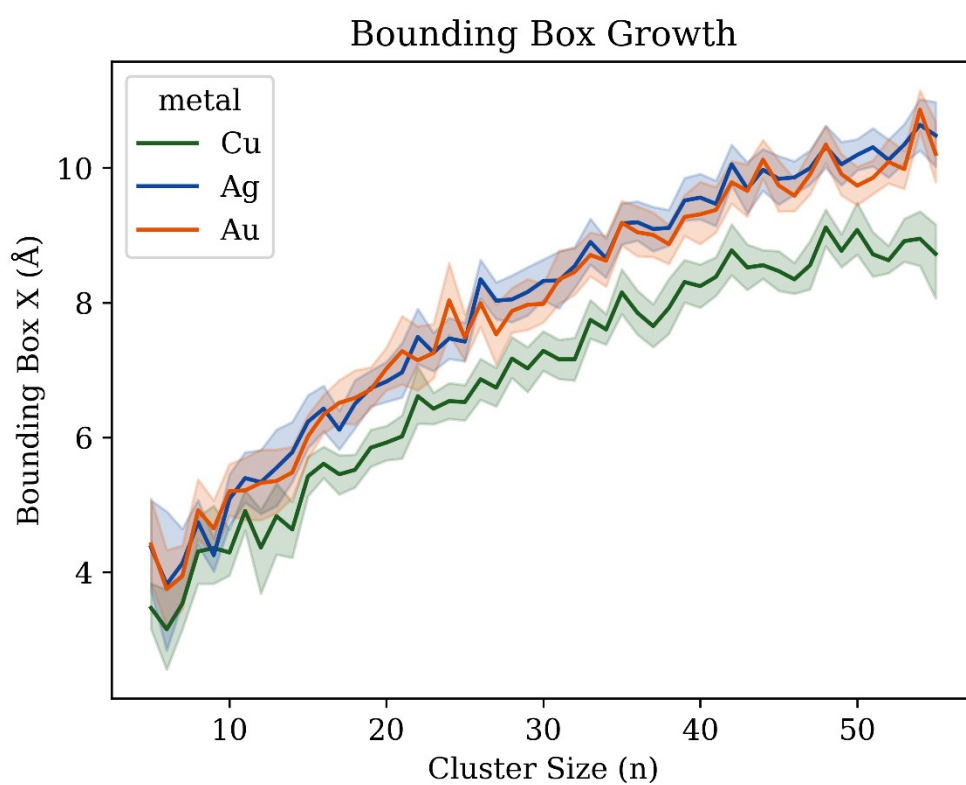
**Figure S3.** Compactness Distribution



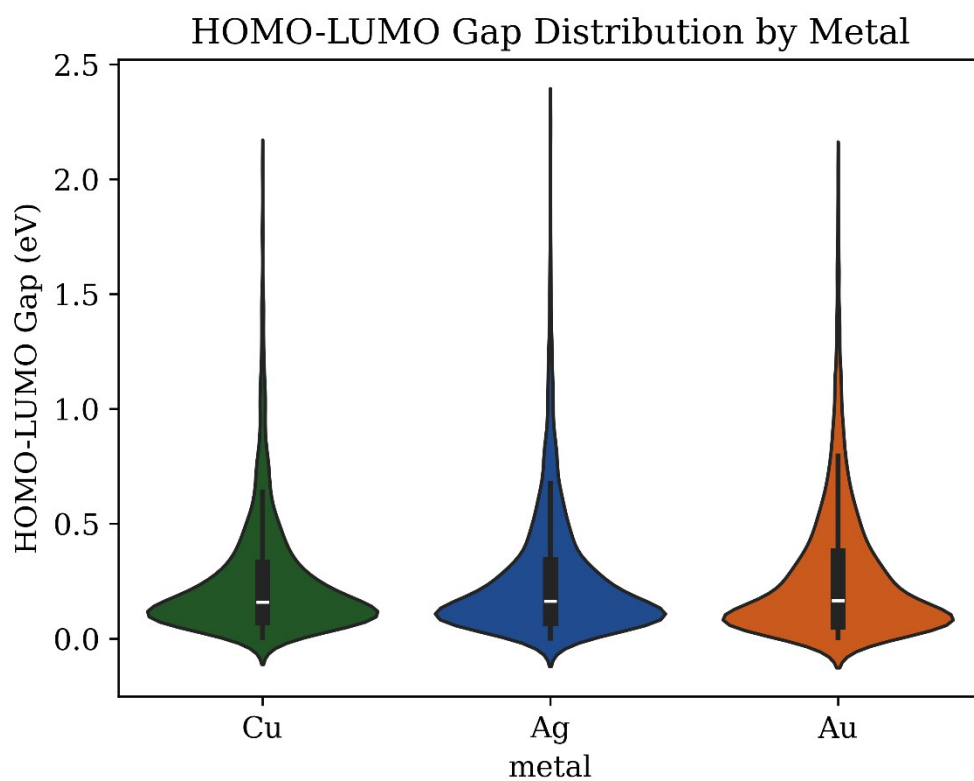
**Figure S4.** Asphericity Distribution By Metal



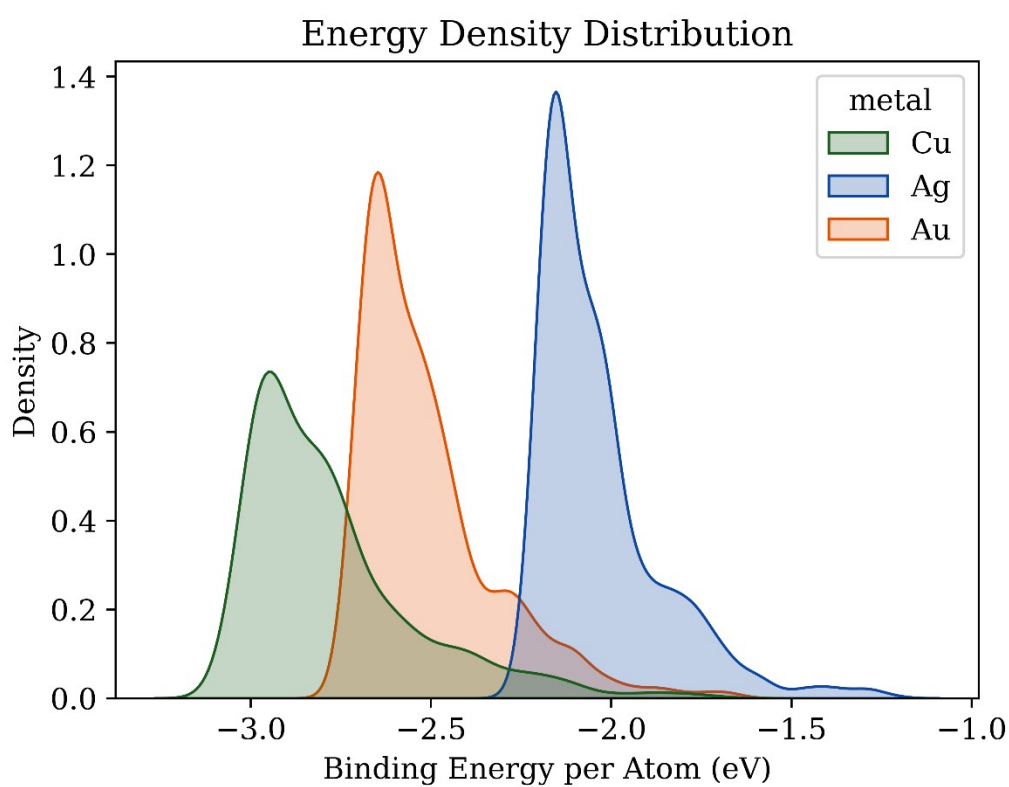
**Figure S5.** Shape Deviation (Asphericity)



**Figure S6.** Bounding Box Growth



**Figure S7.** HOMO-LUMO Gap Distribution



**Figure S8.** Energy Density Distribution

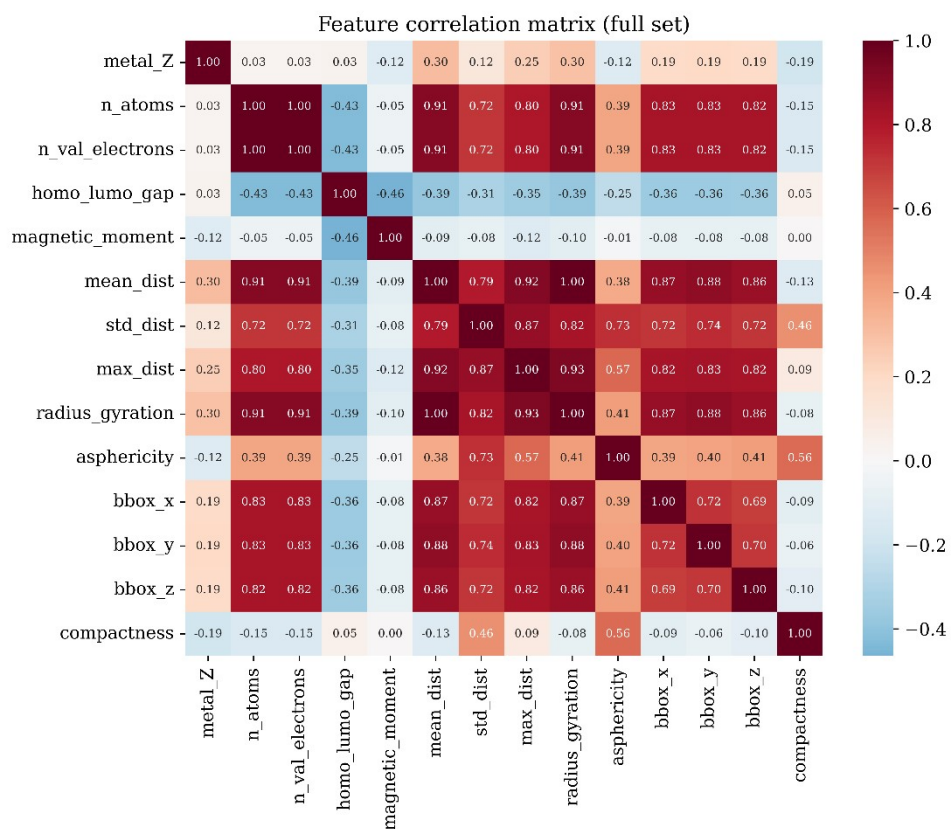


Figure S9. Full Correlation Matrix  
3D Stability Landscape

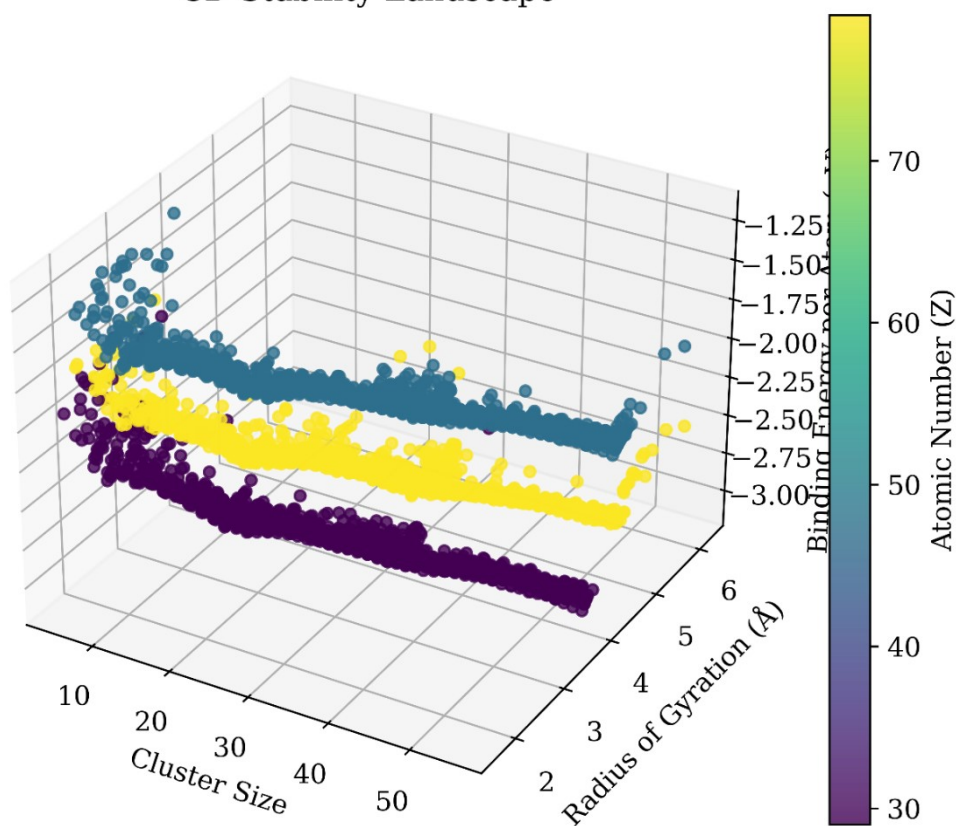


Figure S10. 3D Stability Landscape

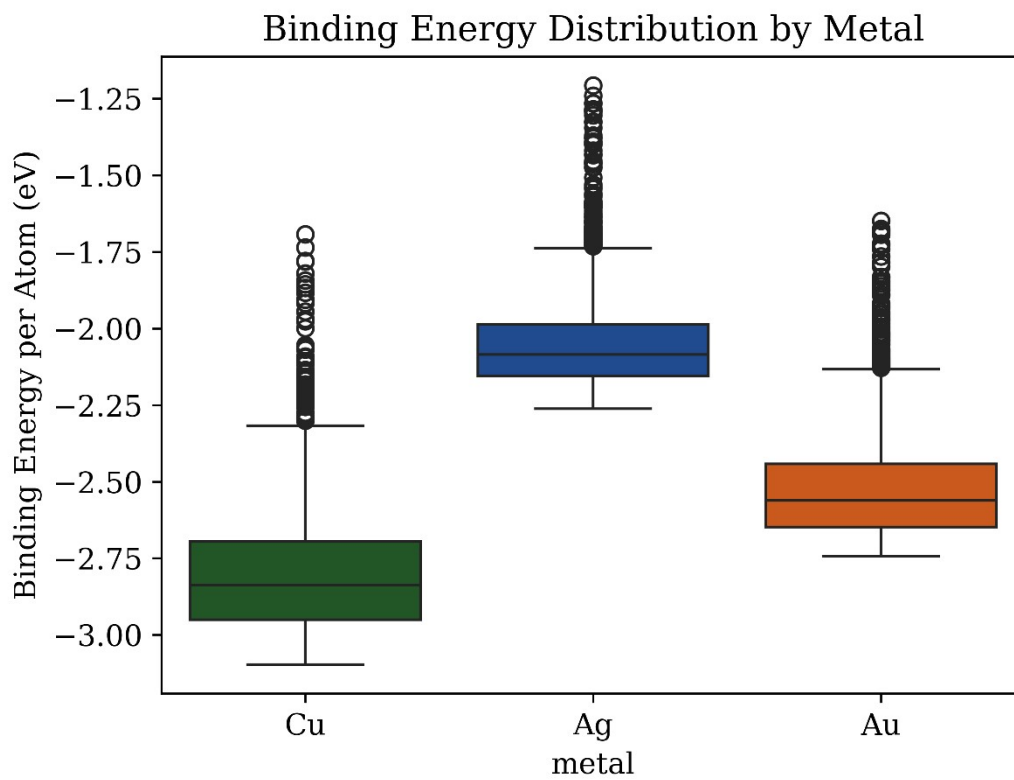


Figure S11. BE by Metal