Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.

Ensemble Machine Learning Methods: Predicting Electron Stopping Powers from Small Experimental Database

[1]Mehnaz, [1]L.H. Yang, [2]B. Da[#], [1]Z.J. Ding[*]

[1]*Hefei National Laboratory for Physical Sciences at Microscale and Department of Physics, University of Science and Technology of China, Hefei, Anhui 230026, P.R. China;*

[2]*Research and Services Division of Materials Data and Integrated System (MaDIS), National Institute for Materials Science (NIMS), 1-2-1 Sengen, Tsukuba, Ibaraki 305-0047, Japan*

[#]e-mail: DA.Bo@nims.go.jp

[*]e-mail: zjding@ustc.edu.cn

Supplementary Information

Table S1: Feature importance for tree-based models (RF, DT, GBR) in different energy ranges. In the first column *E* denotes the entire energy range in which the experimental data is present.

Features	Feature importance for RF			
	E	< 1 keV	1 - 5 keV	> 5 keV
Z	0.15	0.10	0.10	0.03
N_s	0.01	0.03	0.03	0.00
N_p	0.04	0.03	0.03	0.03
N_d	0.01	0.03	0.03	0.03
N_f	0.01	0.03	0.02	0.04
N_v	0.12	0.16	0.11	0.12
N_c	0.12	0.13	0.10	0.09
P	0.11	0.13	0.09	0.03
G	0.04	0.03	0.01	0.02
n	0.01	0.02	0.04	0.03
ρ	0.5	0.46	0.53	0.47
A	0.17	0.14	0.18	0.19
R	0.61	0.51	0.60	0.45
ρZ	0.43	0.31	0.39	0.38
iEA	0.35	0.33	0.41	0.39

Features	Feature Importance for DT			
	E	**< 1 keV**	**1 - 5 keV**	**> 5 keV**
Z	0.14	0.24	0.15	0.12
N_s	0.03	0.03	0.02	0.03
N_p	0.01	0.01	0.01	0.01
N_d	0.08	0.09	0.08	0.08
N_f	0	0.00	0.00	0.00
N_v	0.12	0.13	0.11	0.16
N_c	0.01	0.01	0.01	0.01
P	0.98	0.73	0.81	0.80
G	0.32	0.23	0.33	0.35
n	0	0.00	0.00	0.00
ρ	0.53	0.39	0.42	0.55
A	0.24	0.25	0.25	0.25
R	0.23	0.38	0.32	0.43
ρZ	0.11	0.10	0.10	0.07
iEA	0.14	0.30	0.37	0.28
Features	**Feature Importance for GBR**			
	E	**< 1 keV**	**1 - 5 keV**	**> 5 keV**
Z	0.16	0.13	0.18	0.11
N_s	0.01	0.01	0.02	0.03
N_p	0.09	0.10	0.08	0.11
N_d	0.00	0.00	0.00	0.01
N_f	0.00	0.00	0.01	0.01
N_v	0.01	0.02	0.01	0.02
N_c	0.36	0.36	0.32	0.42
P	0.03	0.05	0.07	0.06
G	0.03	0.05	0.03	0.04
n	0.01	0.01	0.02	0.02
ρ	0.48	0.45	0.46	0.52
A	0.23	0.25	0.23	0.27
R	0.43	0.39	0.44	0.42
ρZ	0.49	0.45	0.48	0.45
iEA	0.29	0.30	0.30	0.30

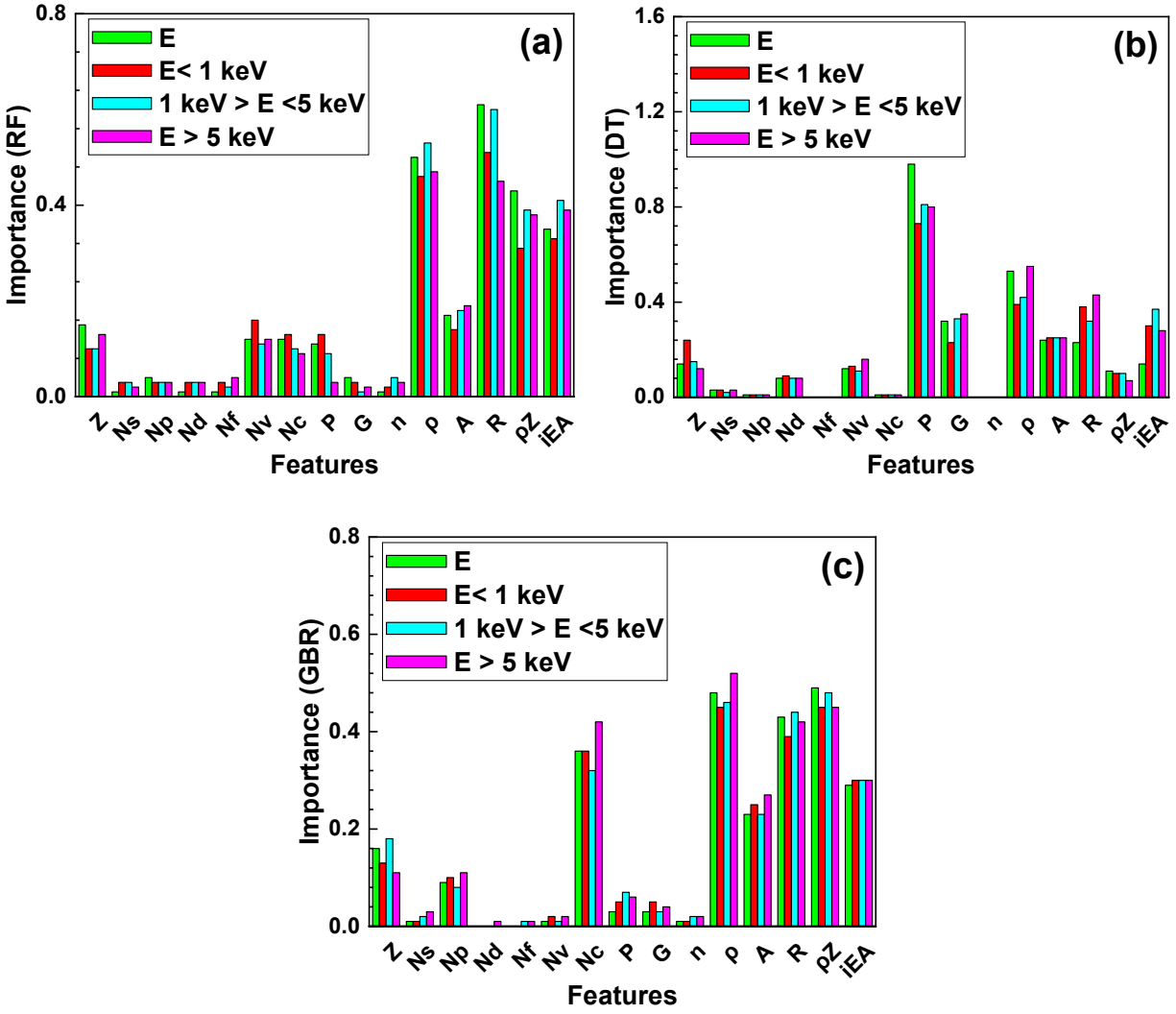

Figure S1: Feature importance for tree-based models over various energy ranges: (a) for RF (b) DTs and (c) GBR. The green color shows the entire energy range for which the experimental data are available.