Electronic Supplementary Material (ESI) for Materials Advances.

Supporting Information for：

Interpretable-machine-learning-guided discovery of dominant intrinsic factors of sensitivity of high explosives

Xianshuang Wang, [ψ] Yage He, [§] Xinyu Zhang, [§] Maoxin Hu, [§] Wanzhu Zhao, [§] Haohan Sun, [§] Xiaoning Yang, [§] Xiaodong Liu and [§] Ruibin Liu, [§,*]

[ψ] School of Materials Science and Engineering, Beijing Institute of Technology, Beijing 100081, China.

[§] Key Lab of Nanophotonics & Ultrafine Optoelectronic Systems, and School of Physics, Beijing Institute of Technology, Beijing 100081, China.

This PDF file includes:

Supplementary Figures

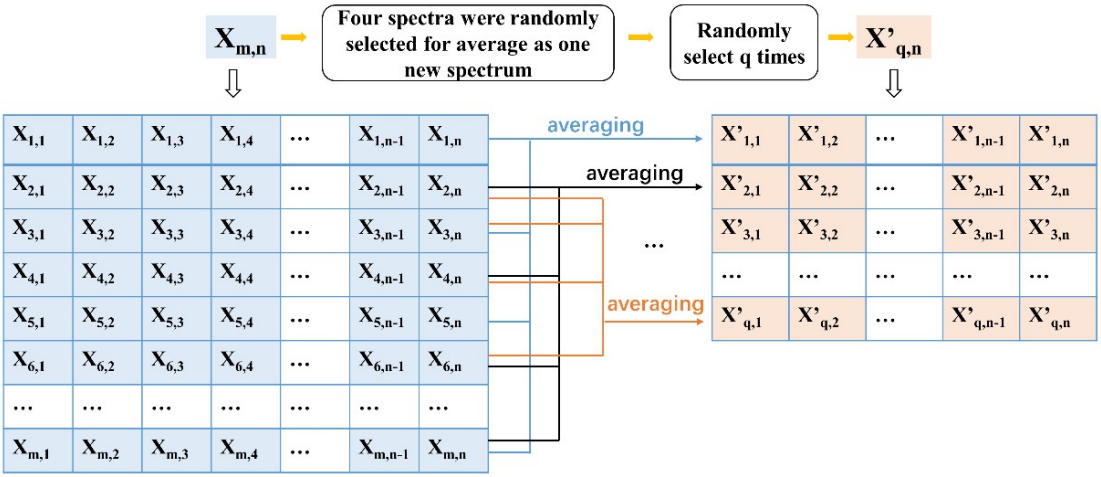

Figure S1. Random sampling and averaging strategy.

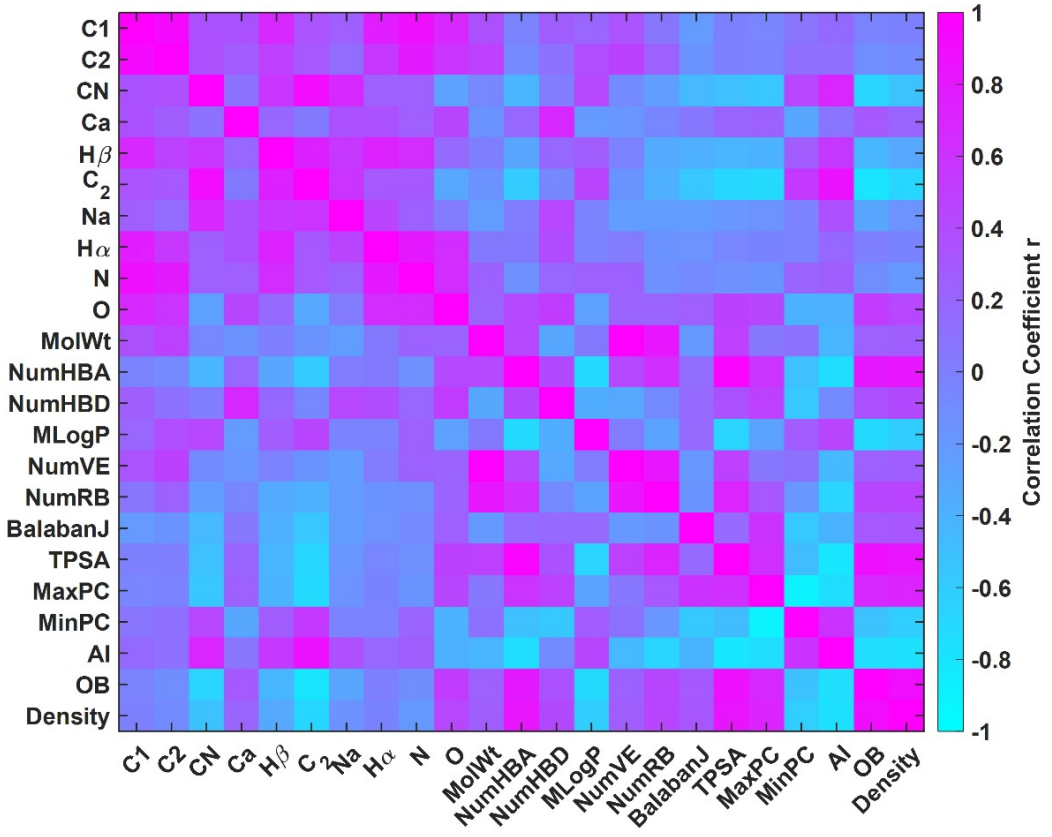

Figure S2. Spearman rank correlation matrices of features.

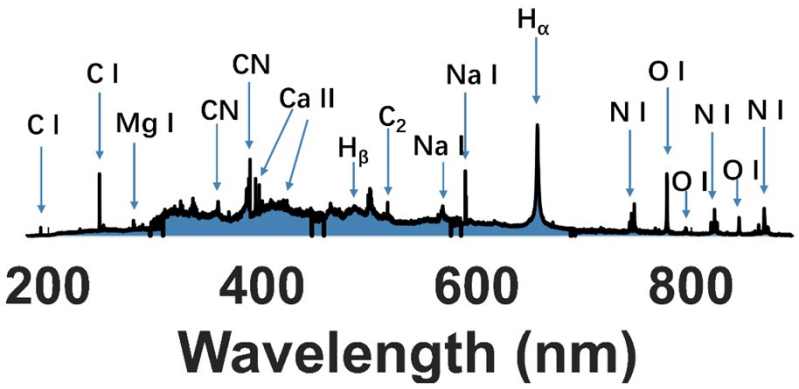

Figure S3. A representative LIPS spectrum of HEs with spectral lines recognition.

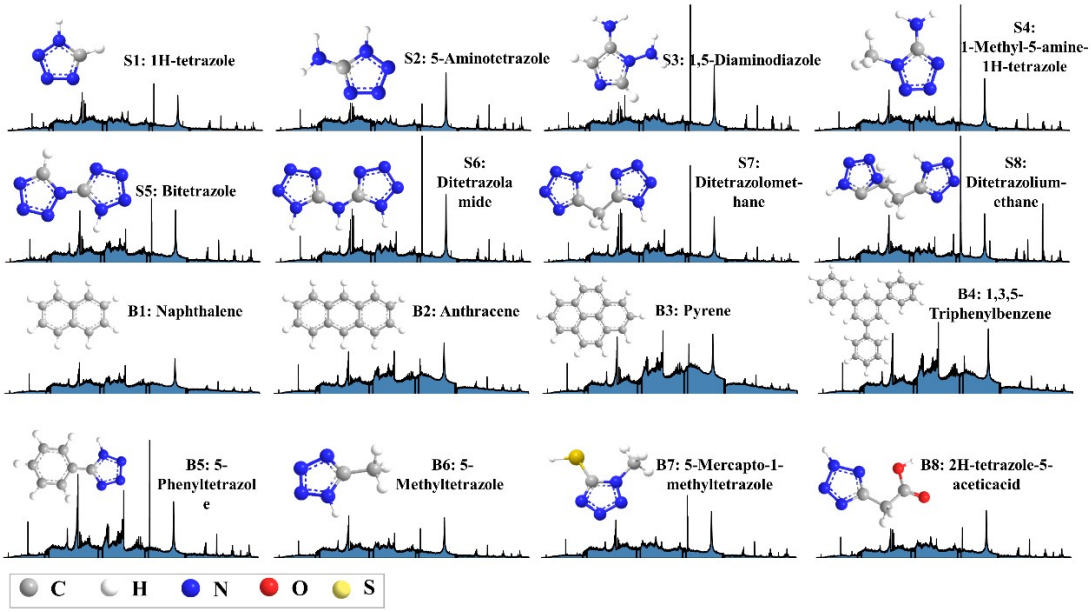

Figure S4. LIPS spectra and molecular structures of eight kinds of energetic tetrazole ring-based organic high-nitrogen compounds and eight kinds of benzene ring organics.

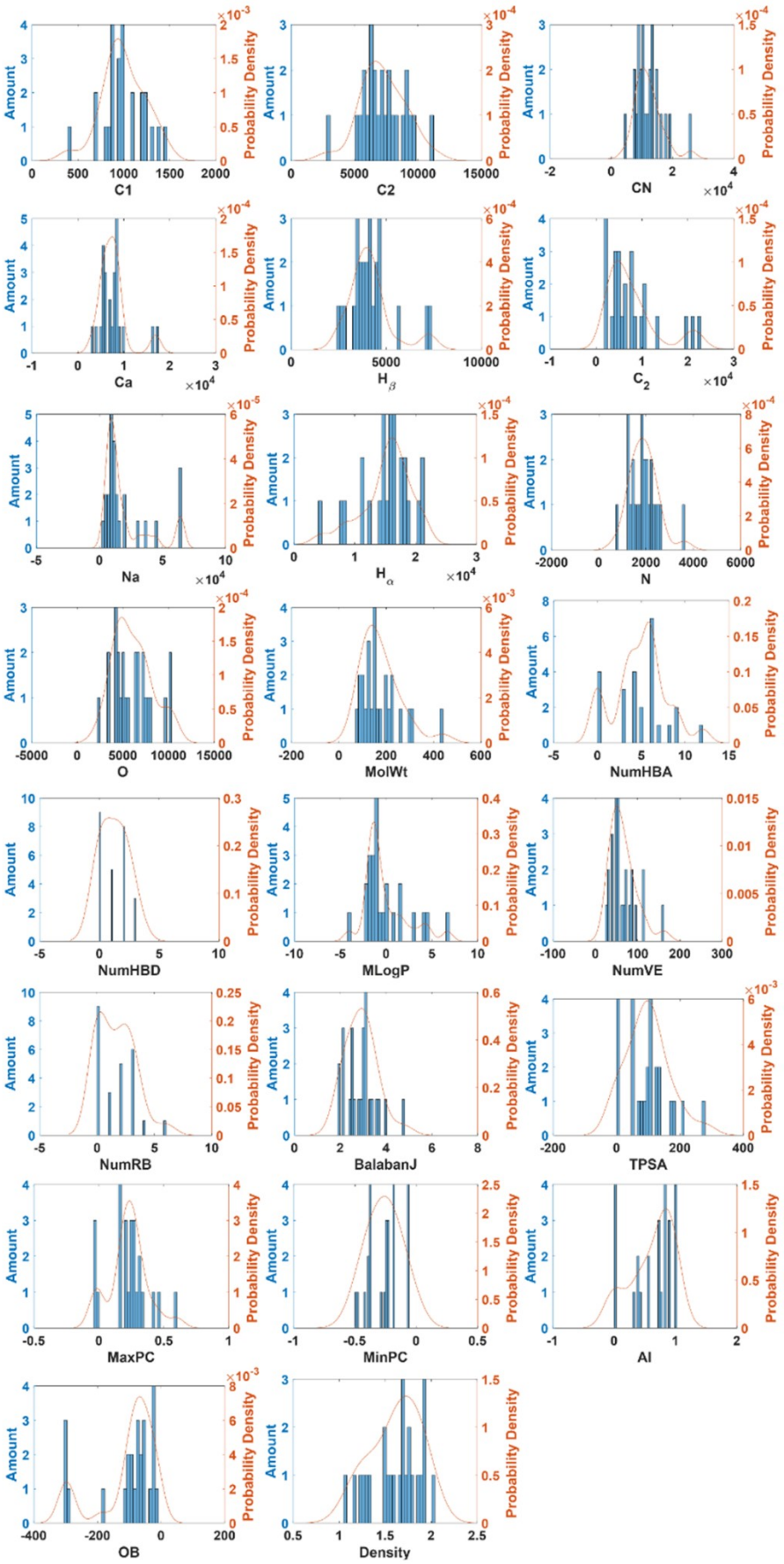

Figure S5. Histograms and probability density distribution of selected features of all data sets.

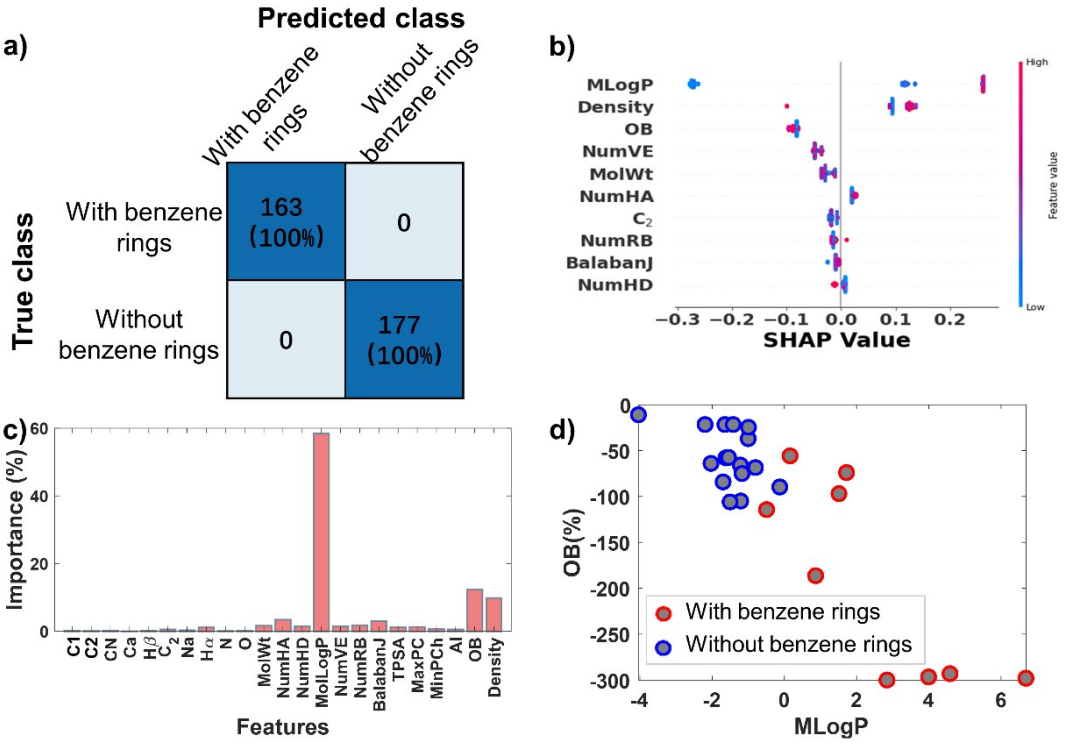

Figure S6. Feature importance analysis of classification model (with benzene rings VS without benzene rings). (a) Confusion matrix figures of prediction results. (b) SHAP analysis of the training data set samples and the RF model. (c) Feature importance directly educed from RF model. (d) Samples distribution in the custom operator feature space of MLogP and OB.

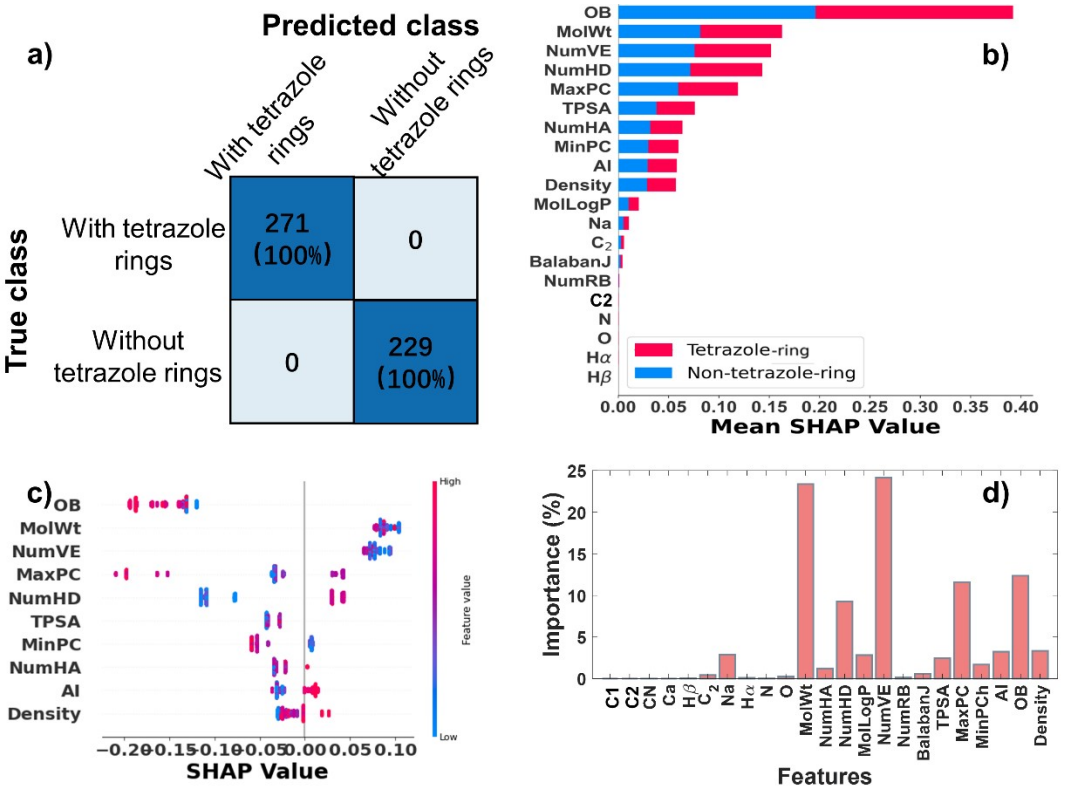

Figure S7. Feature importance analysis of classification model (with tetrazole rings VS without tetrazole rings). (a) Confusion matrix figures of prediction results. (b) Mean SHAP value analysis of the training data set samples and the RF model. (c) SHAP analysis of the training data set samples and the RF model. (d) Feature importance directly educed from RF model.

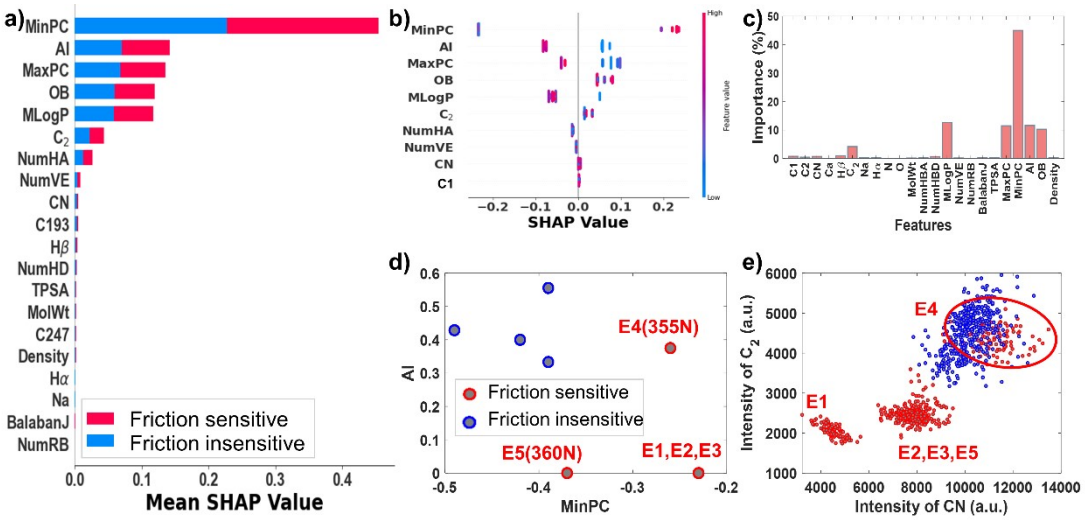

Figure S8. Feature importance analysis of friction sensitivity model. (a) Mean SHAP value analysis of the training data set samples and the RF model. (b) SHAP analysis of the training data set samples and the RF model. (c) Feature importance directly educed from RF model. (d) Samples distribution in the custom operator feature space of MinPC and AI. (e) Samples distribution in the spectral feature space of C_2 and CN.

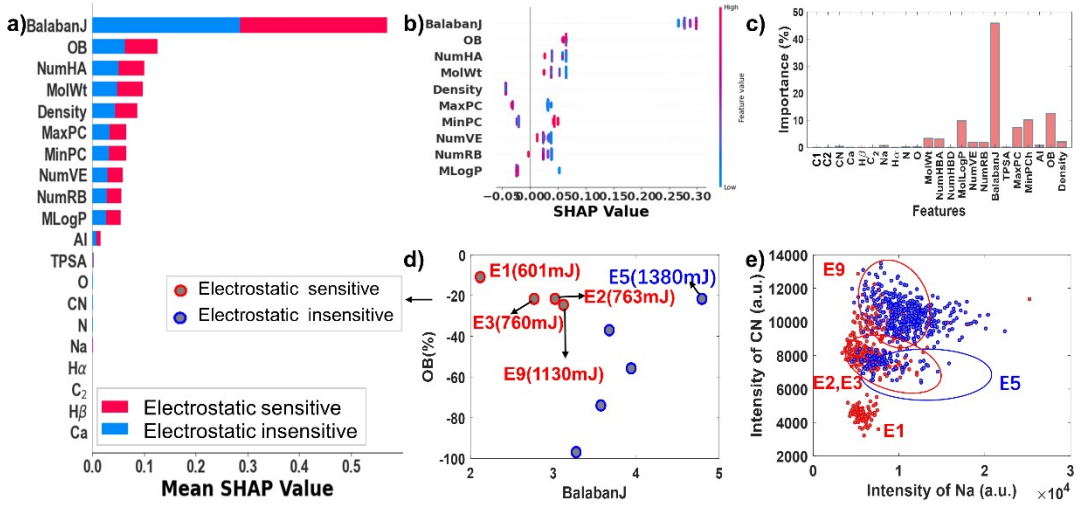

Figure S9. Feature importance analysis of electrostatic sensitivity model. (a) Mean SHAP value analysis of the training data set samples and the RF model. (b) SHAP analysis of the training data set samples and the RF model. (c) Feature importance directly educed from RF model. (d) Samples distribution in the custom operator feature space of BalabanJ and OB. (e) Samples distribution in the spectral feature space of Na and CN.

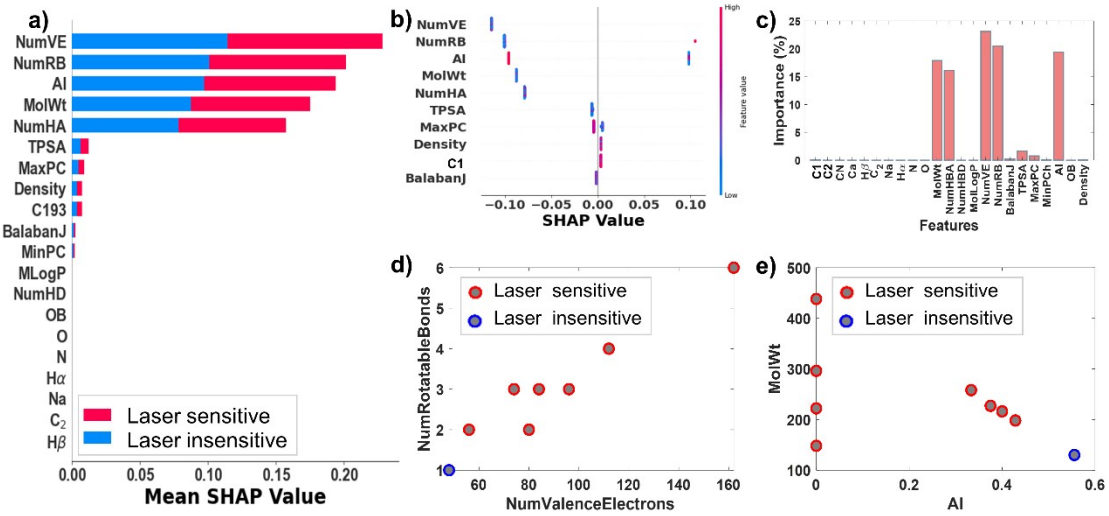

Figure S10. Feature importance analysis of laser sensitivity model. (a) Mean SHAP value analysis of the training data set samples and the RF model. (b) SHAP analysis of the training data set samples and the RF model. (c) Feature importance directly educed from RF model. (d) Samples distribution in the custom operator feature space of NumVE and NumRB (e) Samples distribution in the custom operator feature space of AI and MolWt.

Supplementary Tables

Table S1. A list of organic ring compounds.

Sample	Chemical name	Sample	Chemical name
S1	1H-tetrazole	B1	Naphthalene
S2	5-Aminotetrazole	B2	Anthracene
S3	1,5-Diaminodiazole	B3	Pyrene
S4	1-Methyl-5-amine-1H-tetrazole	B4	1,3,5-Triphenylbenzene
S5	Bitetrazole	B5	5-Phenyltetrazole
S6	Ditetrazolamide	B6	5-Methyltetrazole
S7	Ditetrazolomethane	B7	5-Mercapto-1-methyltetrazole
S8	Ditetrazoliumethane	B8	2H-tetrazole-5-aceticacid

Table S2. Abbreviation and description for features.

Abbreviation	Description	Abbreviation	Description
C1	Spectral intensity of C atom at 193.1 nm	C2	Spectral intensity of C atom at 247.9 nm
CN	Spectral intensity of CN molecule at 388.3 nm	Ca	Spectral intensity of Ca ion at 393.4 nm
H_β	Spectral intensity of H atom at 486.1 nm	C_2	Spectral intensity of C_2 molecule at 516.5 nm
Na	Spectral intensity of Na atom at 589.0 nm	H_α	Spectral intensity of H atom at 656.3 nm
N	Spectral intensity of N atom at 742.4 nm	O	Spectral intensity of O atom at 777.4 nm
MolWt	Molecular weight	NumHBA	Number of hydrogen bond acceptor
NumHBD	Number of hydrogen bond donor	MLogP	Multi-core oil-water partition coefficient
NumVE	Number of valence electrons	NumRB	Number of rotatable bond
BalabanJ	A useful tool for expressing "topological shape" of molecules or of molecular fragments	TPSA	Topological polar surface area
MaxPC	Maximum value of partial charge	MinPC	Minimum value of partial charge
AI	Aromatic index	OB	Oxygen balance
Density	Crystal density	-	-

Table S3. Different types of features

Feature type		Feature Abbreviation
Spectral features by experiments	Atomic spectral features	C1, C2, Ca, H_{α}, Na, H_{β}, N, O
	Molecular spectral features	CN, C_2
Atomic and molecular custom descriptors by calculation.	Atomic custom descriptors	MaxPC, MinPC
	Molecular custom descriptors	MolWt, NumHBA, NumHBD, MLogP, NumVE, NumRB, BalabanJ, TPSA, AI, OB, Density

Table S4. Data set division of multiple sensitivities classification tasks.

Category name	Classification standard	Samples	The number of final spectra of each sample
Impact sensitive	Impact sensitivity≤40 J	E1,E2,E3,E5,E7,E9	50
Impact insensitive	Impact sensitivity>40 J	E4,E6,E8	100
Friction sensitive	Friction sensitivity≤360 N	E1,E2,E3,E4,E5	100
Friction insensitive	Friction sensitivity>360 N	E6,E7,E8,E9	100
Electrostatic sensitive	Electrostatic sensitivity≤1200 mJ	E1,E2,E3,E9	100
Electrostatic insensitive	Electrostatic sensitivity>1200 mJ	E4,E5,E6,E7,E8	100
Laser sensitive	Laser sensitivity≤70 J cm^{-2}	E1,E2,E3,E4,E5,E6,E7,E8	50
Laser insensitive	Laser sensitivity>70 J cm^{-2}	E9	400

Table S5. Data set division of classification tasks for comparison.

Category name	Classification standard	Samples	The number of final spectra of each sample
with benzene rings	Number of benzene rings in molecular structure≥1	E4,E6,E8,B1,B2,B3,B4,B5,B6	100
without benzene rings	Number of benzene rings in molecular structure=0	S1,S2,S3,S4,S5,S6,S7,S8,E1,E2, E3,E5,E7,E9,B7,B8	50
with tetrazole rings	Number of tetrazole rings in molecular structure≥1	S1,S2,S3,S4,S5,S6,S7,S8,B5,B6, B7,B8	100
without tetrazole rings	Number of tetrazole rings in molecular structure=0	E1,E2,E3,E4,E5,E6,E7,E8,E9,B1, B2,B3,B4	100