MXgap: A MXene Learning Tool for Bandgap Prediction

Diego Ontiveros,[1] Sergi Vela,[2] Francesc Viñes,[1] Carme Sousa[1]

[1] *Departament de Ciència de Materials i Química Física & Institut de Química Teòrica i Computacional (IQTCUB), Universitat de Barcelona*

[2] *Institut de Química Avançada de Catalunya (IQAC-CSIC)*

MXenes ($M_{n+1}X_nT_x$)

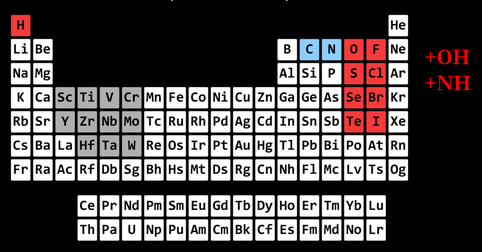
M = Transition Metal (Groups 3 – 6)
X = C or N n = 1 – 4
T = Termination (Groups 16 – 17)

Promising candidates for photocatalysing the water splitting process and produce clean H_2.[1,2]

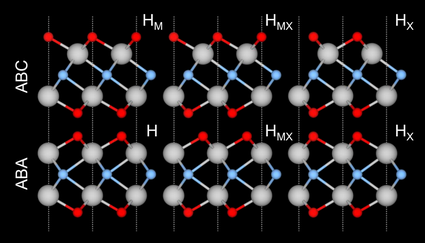
Objective: Train ML models to quickly predict the MXene bandgap.
&
Develop a free open-source, user-friendly Python package (`MXgap`).

Database

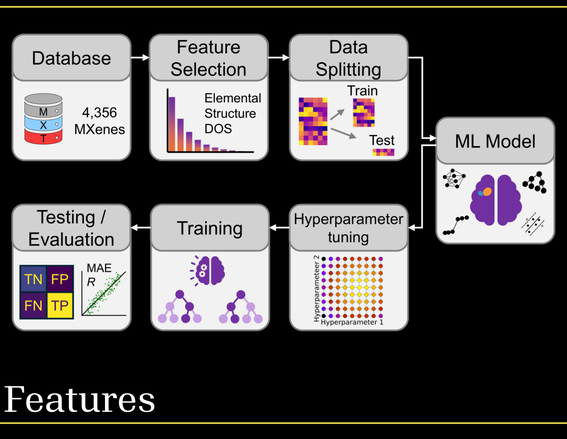
- **4356** different MXene structures.
- Gathered from our previous screenings.[1,2]
- Considered 11 M, C and N, 11 T.



- Also considers n = 1 – 3, with 2 stackings and 3 termination position for each stacking.



Workflow



Features

Initial Training (33 features):
- Elemental: Periodic table properties.
- Structural: Distances, width, etc.

Enhanced Training (136 features):
- Also includes PBE DOS.

Target: Bandgap (PBE0).

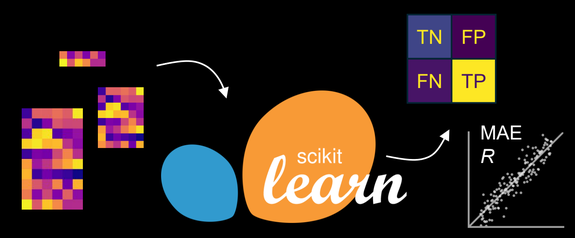
ML Models

Models Used:
- Random Forest (RF)
- Gradient Boosting (GB)
- Multi Layer Perceptron (MLP)
- Support Vector Machine (SVM)
- Kernel Ridge Regressor (KRR)
- Logistic Regressor (LR)

Tasks:
- Classification & Regression

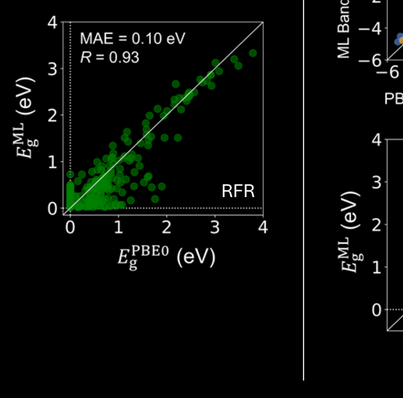
Data Split: 60/20/20 train/validation/test

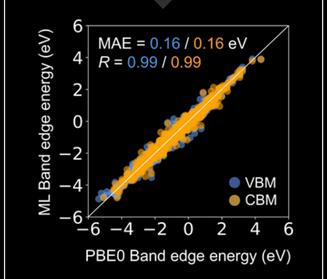
Framework: `Scikit-Learn`

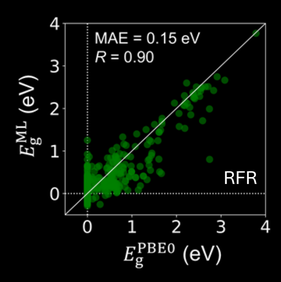


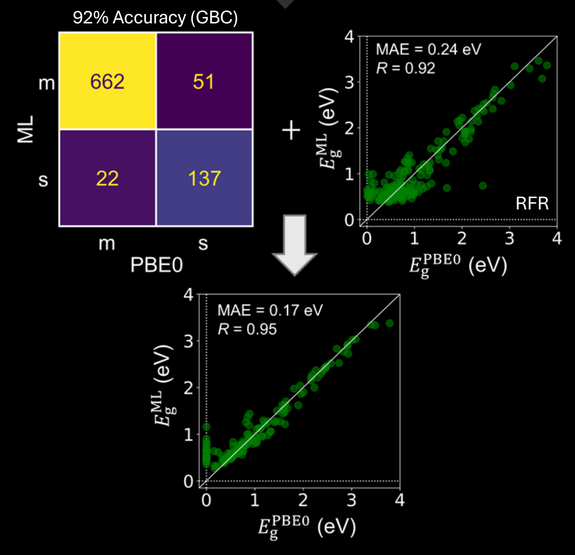
ML Approaches

Full Database	Bandgap Edges	Classifier + Regressor
All MXenes trained together (metallic + semiconductor)	Predict VBM & CBM separately	First, classify MXenes into metallic/semiconductor. Then, predict bandgap for semiconductors only
Low MAE, Good correlation	Excellent VBM/CBM prediction	More accurate results
Misclassification↑ Dispersion↑	Misclassification↑ when combined	Reduced misclassification and dispersion

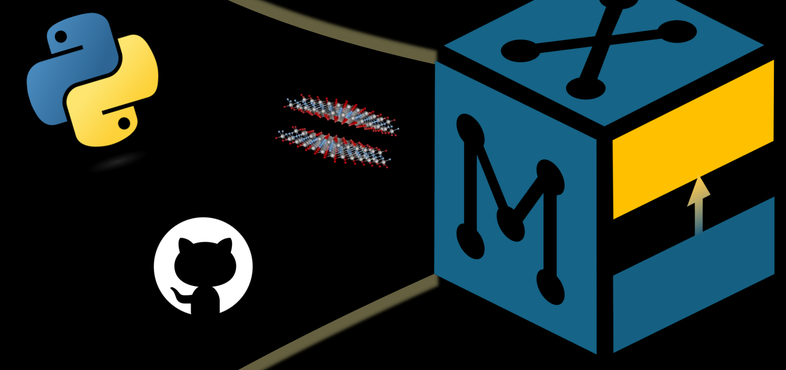








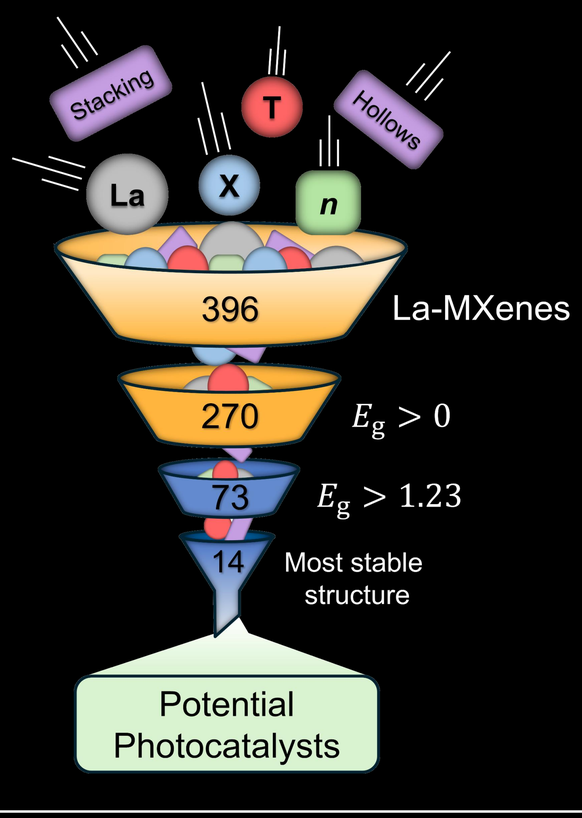
MXgap



- The trained models were integrated into `MXgap`, an open-source Python package developed for predicting MXene bandgaps using ML.
- **How it works?** Automatically parses calculation outputs to obtain relevant features and uses the trained ML models to predict the MXene bandgap.
- **Where?** Free on GitHub and PyPI with full documentation and examples.

Check it out on GitHub!

La-MXenes Screening



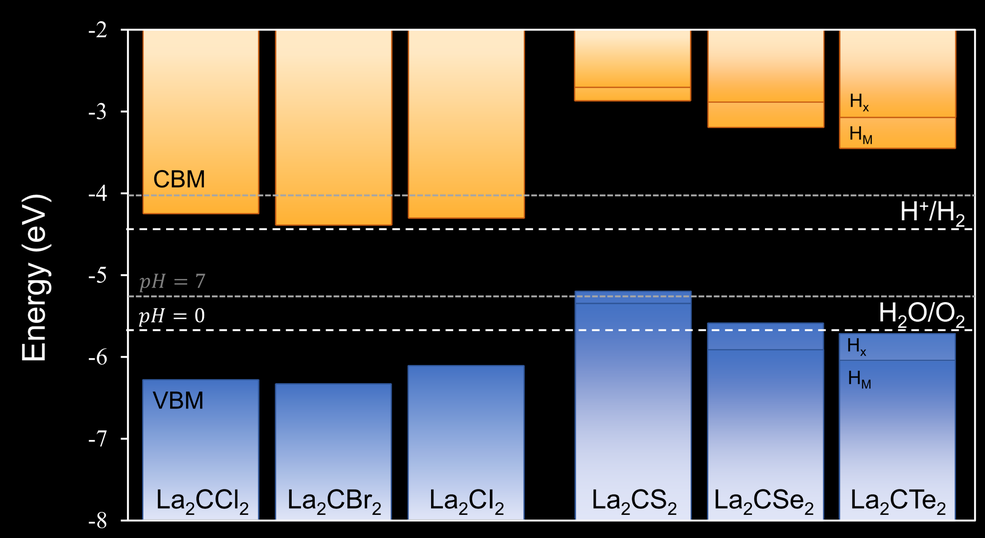
- 396 novel La-MXenes screened with `MXgap` → identifying 14 promising candidates based on their bandgap and stability.
- To confirm these predictions, we selected these 14, plus 14 random MXenes, for PBE0 level validation:
  - Classifier: 89% Accuracy
  - Regression: MAE = 0.25 eV, R = 0.91

Band Alignment

- Suitable band alignment for 6 cases: La_2CT_2 (T = Cl, Br, I, S, Se, Te)
- *pH* = 0: ✓ band alignment for all cases, excepting OER in La_2CS_2.
- *pH* = 7: Some halide-terminated MXenes become unable to photocatalyze HER.

Optical Absorption

- Calculated at GW-BSE level



Optical properties were also calculated:
- Light Absorption: Most MXenes → strong in visible range, especially chalcogen-terminated.
- Solar-to-hydrogen (STH) efficiency: ~40% for halide-terminated *vs.* ~10% for chalcogen.

CONCLUSIONS ✓

Using a ML pipeline combining classification and regression models, MXene bandgaps were accurately predicted. The open-source `MXgap` tool makes these models readily accessible, allowing researchers to rapidly identify promising semiconducting MXenes for photocatalytic applications. The program, applied to screening 396 La-based MXenes, led to the discovery of 6 candidates with suitable band alignments, underscoring the potential of ML to accelerate the research of functional 2D materials for clean energy technologies.

REFERENCES

[1] D. Ontiveros, F. Viñes, C. Sousa, *J. Mater. Chem. A,* 2023, 11, 13754–13764.

[2] D. Ontiveros, S. Vela, F. Viñes, C. Sousa, *Energy Environ. Mater.*, 2024, 7, e12774.

github.com/diegonti/mxgap

diego.ontiveros@ub.edu

diegonti.github.io

Acknowledgments

financiado por MICIU/AEI/10PID2021-126076NB-I00.13039/501100011033 y por FEDER,UE

QHS-2023-2-0017
QHS-2023-3-0012

CEX2021-001202-M Ac2216561 2021SGR00079