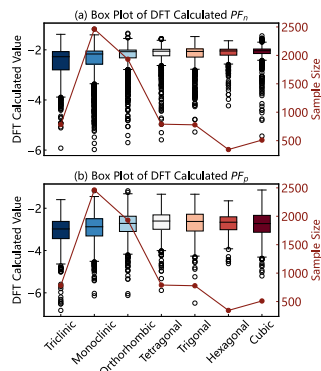


"boltztrap_mp" Dataset

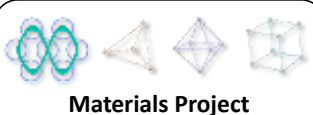


Preprocessing:

- Normalization.
- Outlier del.
- Missing value.
- Log. trans.

$$PF' = \log(PF)$$

Feature Construction:

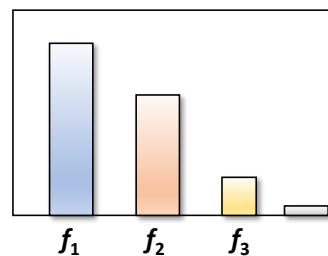
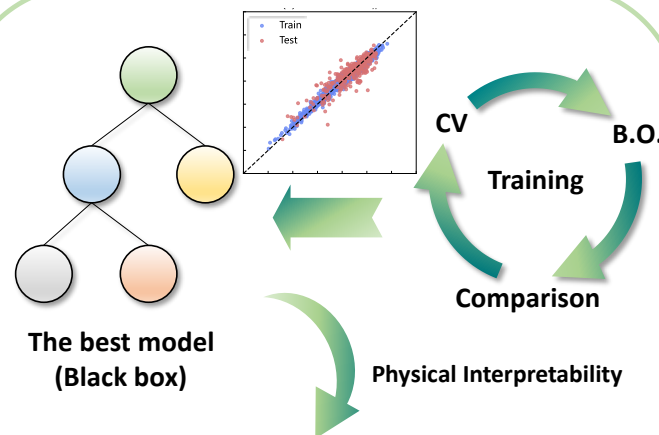


- Energy above hull
- Formation energy
- Crystal system
- Space group number



- ElementFraction
- Stoichiometry
- BandCenter
- IonProperty
- ElementProperty
- ValenceOrbital

ML Training



Feature selection (FIR & R)

$$D_1 = F_1(f_1, f_2, f_3, \dots)$$

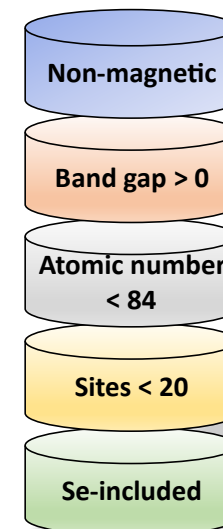
$$+ D_2 = F_2(f_1, f_2, f_3, \dots)$$

$$D_3 = F_3(f_1, f_2, f_3, \dots)$$

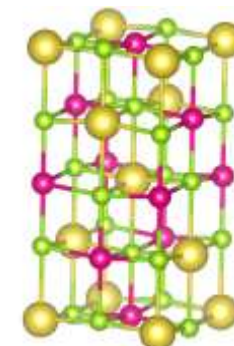
Symbolic Regression (SISSO)

DFT Validation

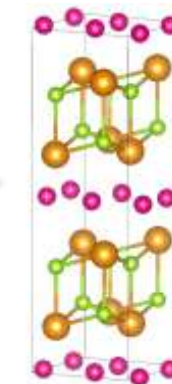
New samples from MP



Validated by
DFT



NaAsSe₂, p-type



NdAsSe, n-type