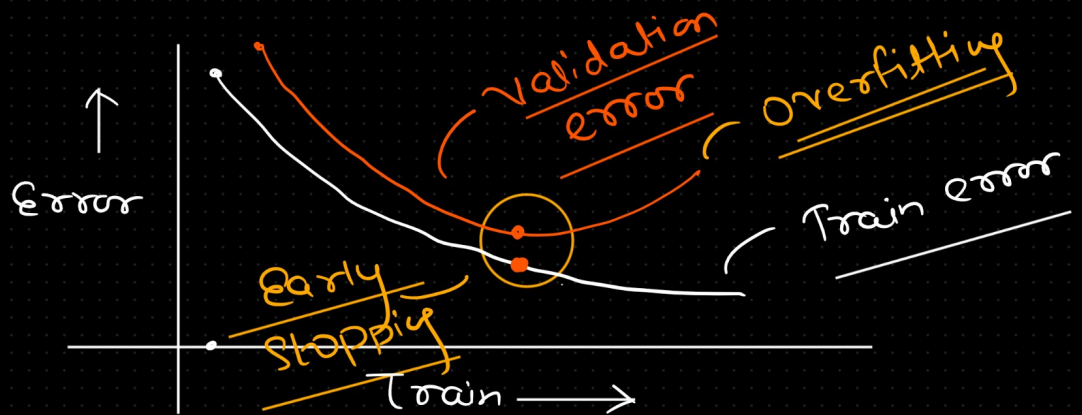


Xg Boost

↳ Distributed Computing → Large dataset
(Big Data) (Parallel
Speed → Computation)

↳ Early stopping → Prevent Overfitting
↳ validation ↓ ↓

train
error ↓ ↓



③ Handle missing data (Best imputation)

④ Xgb Tree (Pruning)
* ↳ avoid overfitting

⑤ Regularization Parameter
* ↳ λ

⑥ Compatibility

→ Really Powerful → Real-time Projects
OF
Kaggle/Hackathon