

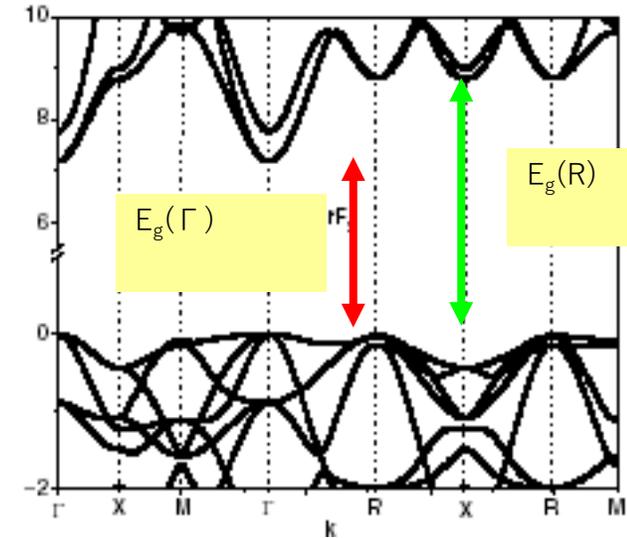
課題ID「jh190019-NAJ」

全電子混合基底第一原理計算法を活用した ネットワーク型エネルギー絶対値算定 マテリアルズ・インフォマティクス

川添良幸（東北大学 未来科学技術共同研究センター）
水関博志(KIST)、大野かおる(横国)、佐原亮二(NIMS)、
本郷研太(JAIST)、南里豪志(九大)

本研究の目的と概要

- 発光素子用材料の理論的探索
 - III-V族化合物半導体系のバンドギャップ・エンジニアリング
 - 「計算科学的網羅的探索」から「データ科学的網羅的探索」へ
 - 探索空間拡張→未知の新奇材料発見
 - マテリアルズ・インフォマティクス (MI) 技法の適用
 - 探索空間の効率的探索
 - 標準的なDFTでデータ生成→バイアスされたデータ空間
- 高精度・高効率MI技法の確立
 - 信頼性の高いデータ空間生成：高精度第一原理計算（全電子GW）
 - TOMBO (Tohoku Mixed-Basis Orbitals *Ab Initio* Simulation Package)
 - 機械学習による高精度物性予測
 - 現象理解/効率的探索のための記述子解析
 - MIプラットフォームの構築



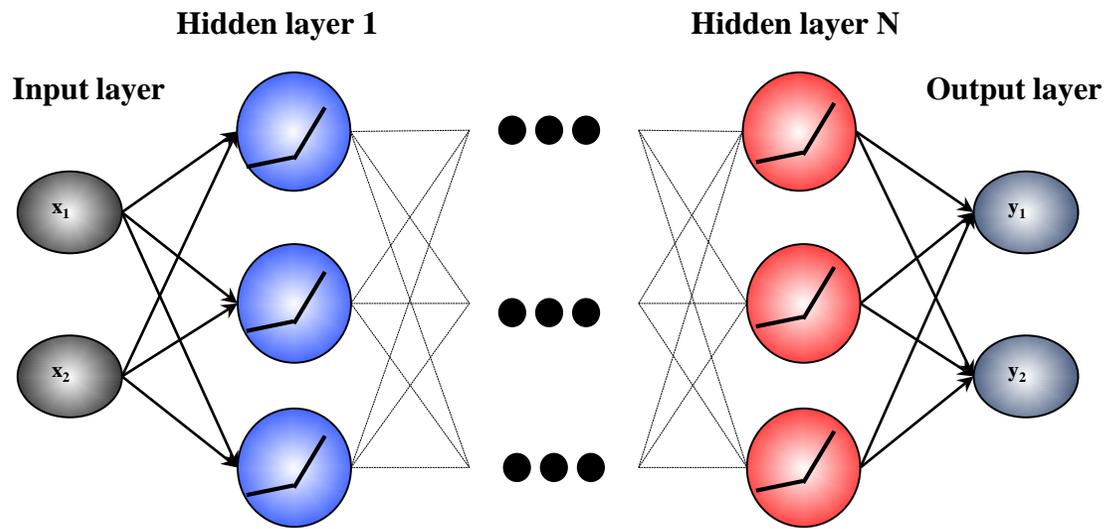
ニューラルネットワーク (NN) に対する入力データと教師データ

0.793733
0
0
1
7
8
0
0
0
1.030930
0
1
1
6
8
0
0
0
1.390473
0
1
2
5
8
0
0
0
...
...
...

バンドギャップ値 (eV)
または
ZBとWZ間のエネルギー差 (eV)

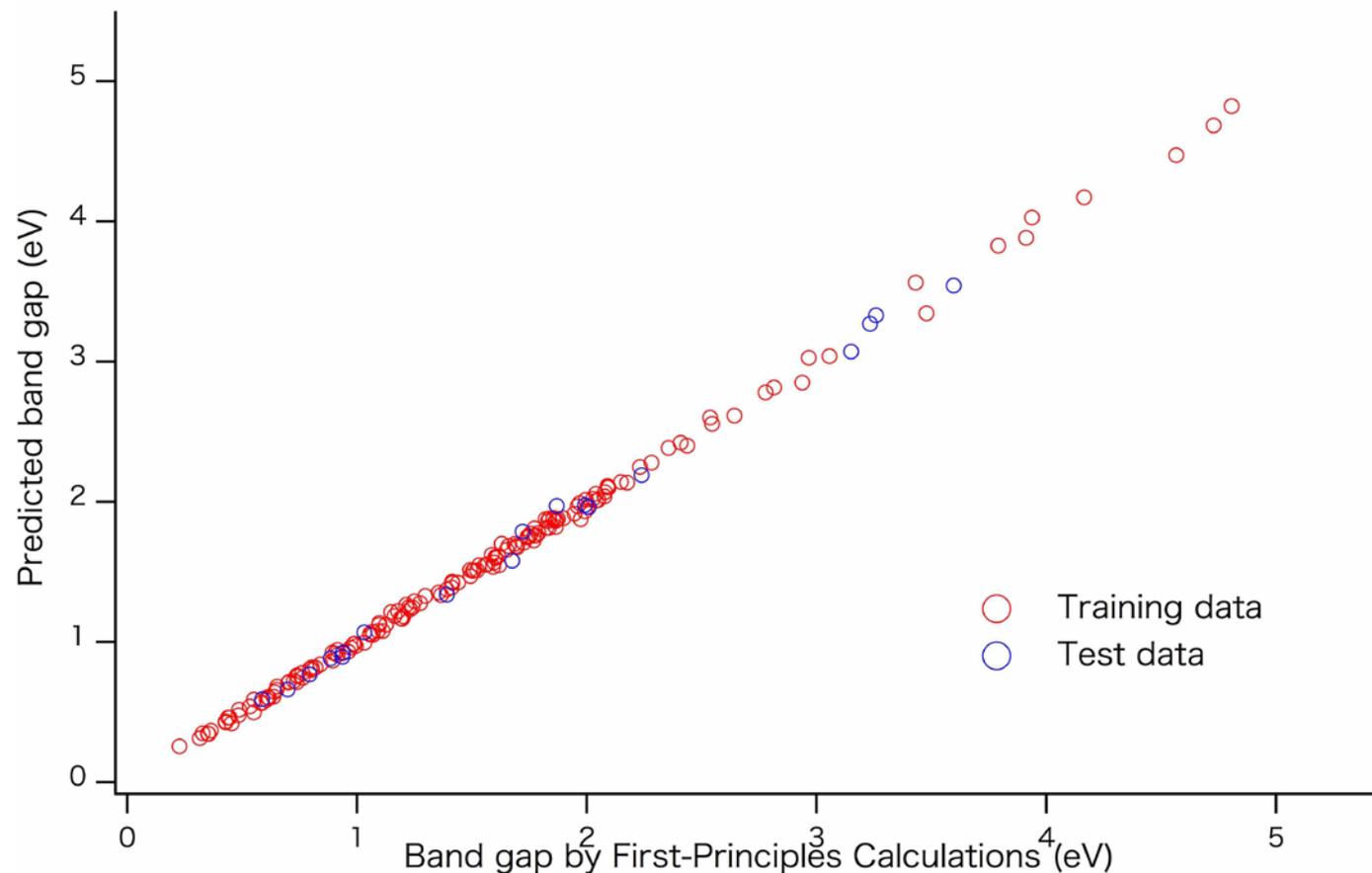
of B
of Al
of Ga
of In
of N
of P
of As
of Sb

教師データ
(第一原理計算によるバン
ドギャップと全エネルギー値)



トレーニングデータ・セット数: 162
テストデータ・セット数: 18

予測されたバンドギャップ値 (Zincblende構造)

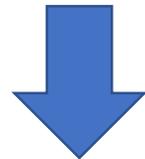


mBJ汎関数を用いたDFT計算によるバンドギャップ値
(PBE汎関数計算による構造最適化)
 $\text{Al}_{1-x}\text{Ga}_x\text{In}_y$ (N,P, As, Sb)

トレーニングデータセット数: 162
テストデータセット数: 18

ニューラルネットワーク計算でのまとめ

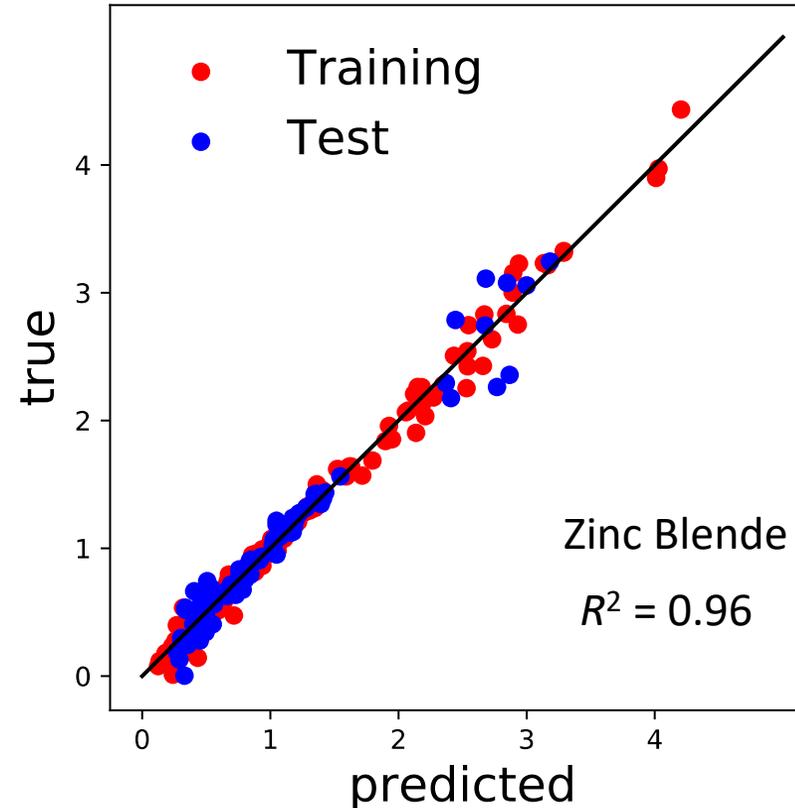
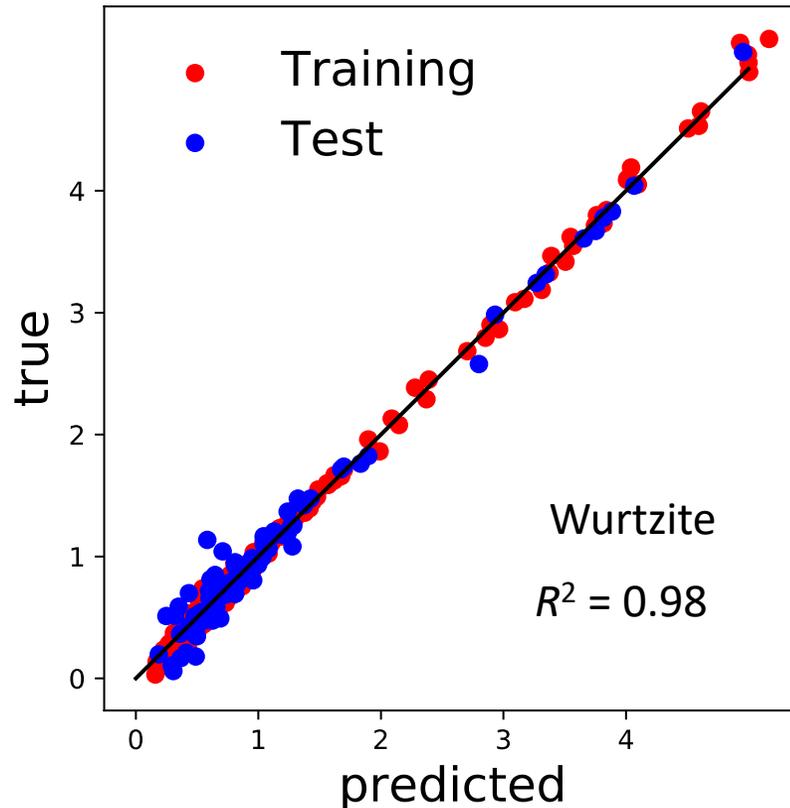
- III-V族半導体の組成情報のみからバンドギャップ値を予想できることを示した。
- 計算を実施していない組成範囲も予測可能になった。
- 効いているパラメータ（元素種）の探索はニューラルネットワークでは困難



マテリアルズ・インフォマティクス手法の適用

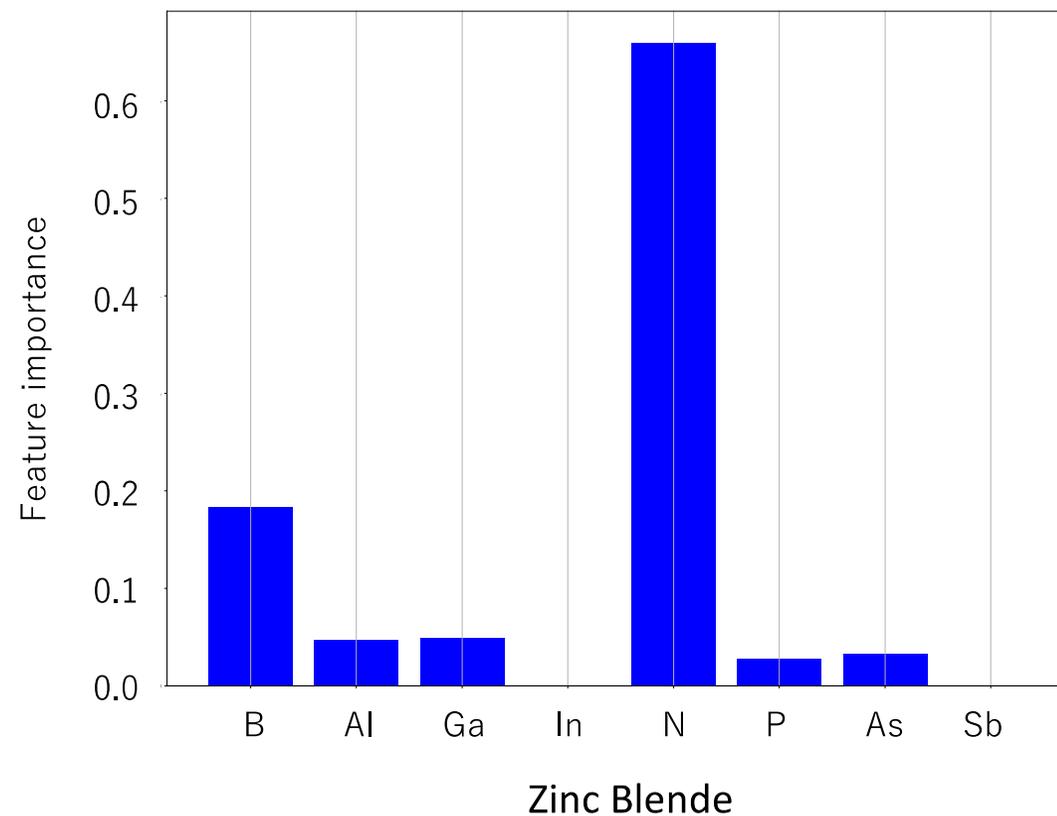
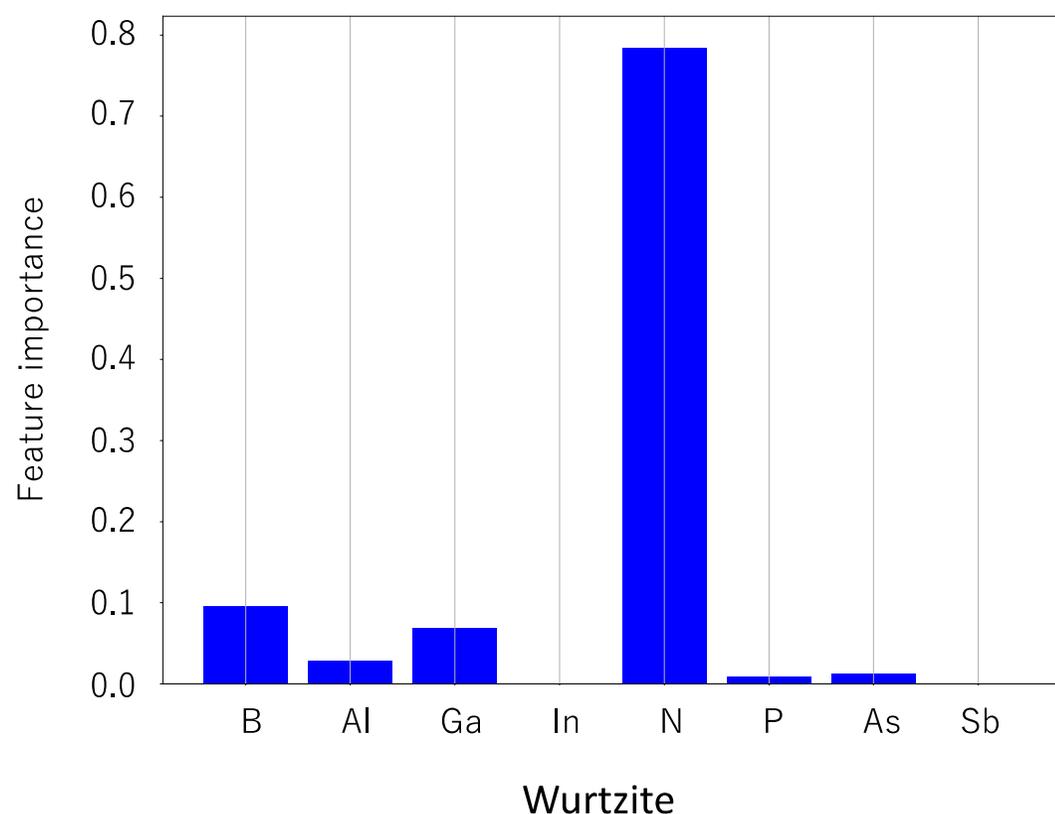
ランダムフォレスト (RF) モデル

- 決定木を弱学習器とするアンサンブル学習
 - 弱学習器による化合物特徴量空間の網羅的カバー：**高汎化性能**
 - 記述子（説明変数）の重要度（寄与度）を算出可能：**解釈可能**（NNでは困難）
 - ブートストラップによるランダムサンプリングで学習データをサンプリング：**少数データ**



RFによる記述子解析

- III族元素(B/Al/Ga/In)とV族元素(N/P/As/Sb)で、Nは一定
- Bの寄与が、WurtziteとZinc Blendeで大きく異なる



- 元素以外の化学情報記述子→物質設計に利用可能な記述子抽出

まとめと展開

- 発光素子としてIII-V族化合物半導体系を対象に網羅的探索
- 要求波長、直接遷移、等の条件を同時に満たす新規材料
- マテリアルズ・インフォマティクス技法適用による高速化
- $\text{Al}_{1-x-y}\text{Ga}_x\text{In}_y$ X系で最適材料探索



- 今後
- DFTからGW計算へ
- 機械学習によるDFT結果のGW計算への調整