SAND2020-0961C

Sandia National Laboratories

Extracting an Empirical Intermetallic Hydride Design Principle from Limited Data via Interpretable Machine Learning

Matthew Witman,[1] Sanliang Ling,[2] David M. Grant,[2] Gavin S. Walker,[2] Sapan Agarwal,[1] Vitalie Stavila,[1] Mark Allendorf[1]

[1]Sandia National Laboratories, Livermore, CA; [2]Advanced Materials Research Group, University of Nottingham, UK

Introduction

An open question in the metal hydride community is whether there are simple, physics-based design rules that dictate the thermodynamic properties of these materials across the variety of structures and chemistry they can exhibit. While black box machine learning (ML)-based algorithms can predict these properties with some success, they do not directly provide the basis on which these predictions are made, therefore complicating the a priori design of novel materials exhibiting a desired property value. In this work we demonstrate how feature importance, as identified by a gradient boosting tree regressor, uncovers the strong dependence of the metal hydride equilibrium H_2 pressure on a volume-based descriptor that can be computed from just the elemental composition of the intermetallic alloy. Elucidation of this simple structure–property relationship is valid across a range of compositions, metal substitutions, and structural classes exhibited by intermetallic hydrides. This permits rational targeting of novel intermetallics for high-pressure hydrogen storage (low-stability hydrides) by their descriptor values, and we predict a known intermetallic to form a low-stability hydride (as confirmed by density functional theory calculations) that has not yet been experimentally investigated [1].

Data cleaning

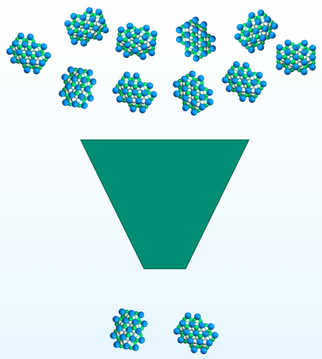
HydPARK is Sandia's historical dataset on metal hydride properties:

Class	Composition	H wt %	ΔH (kJ/molH_2)	T (K)	P_{eq} (atm)
A_2B	$Ce_{0.5}Th_{1.5}Al_1$	0.4	133	650	0.0003
AB_5	$La_{0.75}Ce_{0.25}Ti_{0.1}Cu_{0.9}Ni_4$	1.4	38.4	22	1.1

Target property: Compute equilibrium pressure at room temperature,

$$\ln P^o_{eq} = -\Delta H/RT^o + \Delta S/R$$

to gauge a material's thermodynamic utility in a given storage application



Initial: ~ 2500 HydPARK compositions

1. Remove duplicate compositions
2. Remove entries with missing data, i.e. $\ln P^o_{eq}$ can't be computed

Final: ~ 400 "ML ready" compositions

ML Model Development

1. Magpie presents a useful way to featurize inorganic materials in terms of elemental properties when exact crystal structures are not known [2]

$$\mathbf{x}_{\mathrm{LaNi5}} = [\,\bar{v}^{Magpie}_{pa}, \bar{T}_{melt}, \dots\,] \in \mathbb{R}^{145}$$

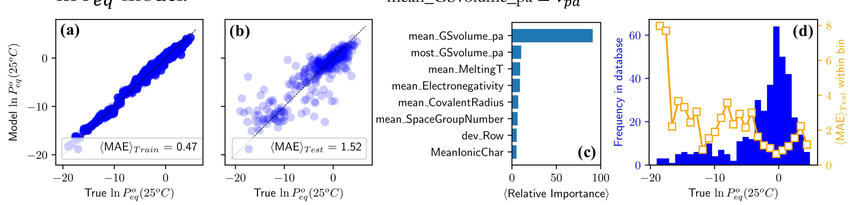
$$\bar{v}^{Magpie}_{pa} = \sum f_i v_i \quad , \quad f_i \equiv \text{composition fraction of element } i$$
$$v_i \equiv \text{ground state volume per atom of solid element } i$$

2. Materials Project (MP) provides electronic structure data if crystal structure known (only ~70 HydPARK compositions found in MP)

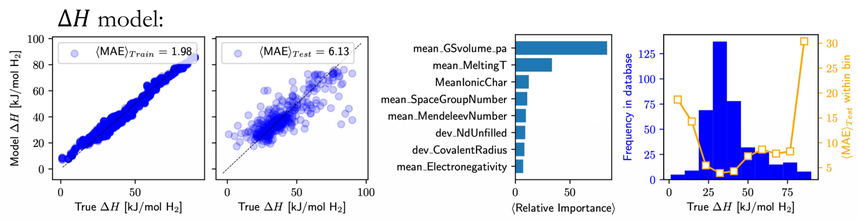
$$v^{MP}_{pa} = V_{cell}/n_{atoms} \qquad V_{cell} \equiv \text{DFT optimized cell volume}$$

3. Gradient Boosting Regressors are "interpretable" : they reveal the relative importance of each feature in the model's prediction [3]

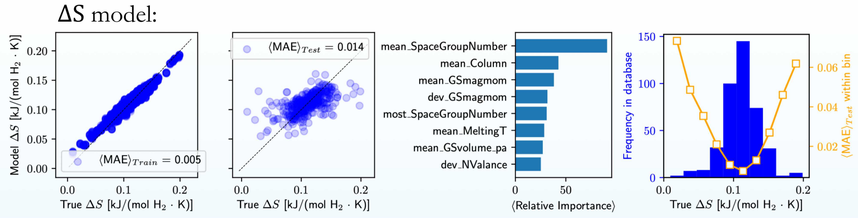
$\ln P^o_{eq}$ model:



ΔH model:

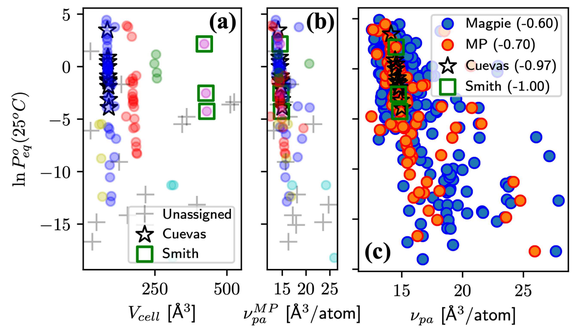


ΔS model:



5-fold Train Performance | 5-fold Test Performance | 5-fold Averaged Feature Importance | Distributed Error

Exploiting Explainable Insights



1. Reveals log-linear relationship with v_{pa}
2. Unifies previous experimental results with a single structure-property relationship
3. Provides an empirical design rule to target low, moderate, or high stability hydrides

Novel Hydride with Targeted Thermodynamics

Utilize DFT to (1) gain insight into the v_{pa} structure-property relationship and (2) validate the hypothesis that a known intermetallic (UNi_5 whose hydriding properties have not yet been tested) would form an ultra low-stability (high-pressure) hydride:

For $ANi_5 \rightarrow ANi_5H_7$ (A=U, Ce, La)

$E_f \equiv$ formation energy of ANi_5H_7

$\Delta E_{def} \equiv$ deformation energy to accommodate H absorption

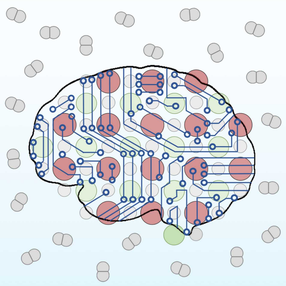
$\Delta E_H \equiv$ binding energy of H

$V/V_0 \equiv$ volume expansion upon H absorption

	v_{pa}	ΔH	E_f	ΔE_{def}	ΔE_H	V/V_0
UNi_5	13.17	-0.60	-285	65.2	-65.8	1.278
$CeNi_5$	13.76	-20.5	-353	49.3	-69.8	1.266
$LaNi_5$	14.38	-36.1	-224	44.3	-80.5	1.256

Conclusions

1. ML models and experimental data provide a powerful tool to explore phenomena too expensive for computational approaches
2. Equilibrium pressure in intermetallic hydrides can be predicted just from the alloy composition, despite noisy/incomplete data
3. Explainable insights from the ML model permit the rational design of novel materials with targeted thermodynamic properties



Sandia National Laboratories is a multimission laboratory managed and operated by National Technology &Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.
SAND No. _____

The authors gratefully acknowledge research support from the U.S. Department of Energy, Office of Energy Efficiency and Renewable Energy, Fuel Cell Technologies Office through the Hydrogen Storage Materials Advanced Research Consortium (HyMARC). This work was supported by the Laboratory Directed Research and Development (LDRD) program at Sandia National Laboratories.

[1] Witman, M. et al. *J. Phys. Chem. Lett.* **2020**, 11, 1, 40-47
[2] Ward, L. et al. *npj Comput. Mater.* **2016**, 2, 16028
[3] Hastie, T. et al. *The Elements of Statistical Learning.* **2009**

Sandia National Laboratories