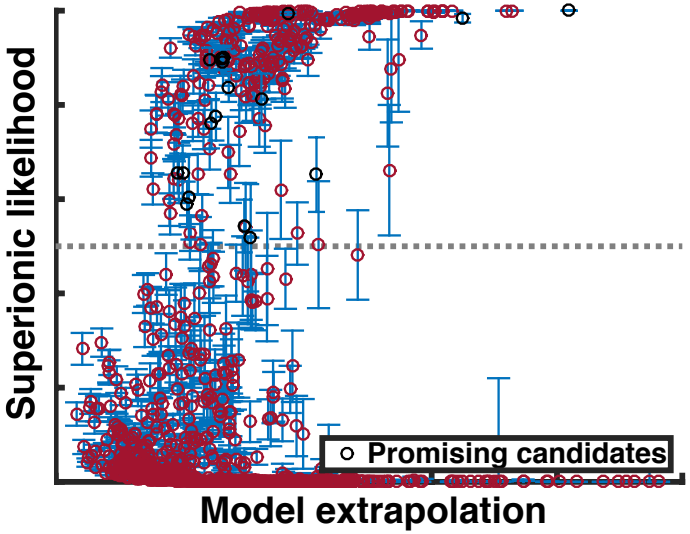
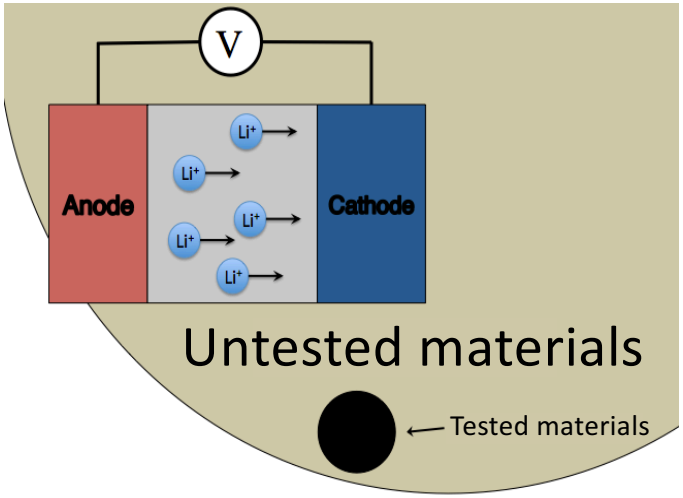
Designing a better battery with machine learning





Austin D. Sendek, Ekin D. Cubuk, Qian Yang, Gowoon Cheon, Evan R. Antoniuk, Karel-Alexander N. Duerloo, Yi Cui, Evan J. Reed

MATLAB Expo 2017

Lithium-ion batteries will enable major innovations in energy

Automotive

Renewable grid

Drones/aviation

Mobile electronics

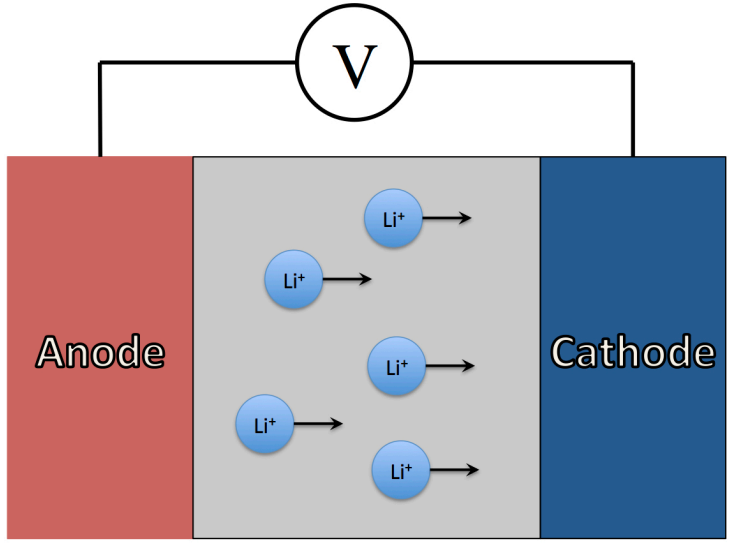
Today’s lithium-ion batteries still face many challenges

- Safety
- Energy density
- Cost
- Cycle life

Renewable grid

Mobile electronics

A major problem worth solving: liquid electrolytes



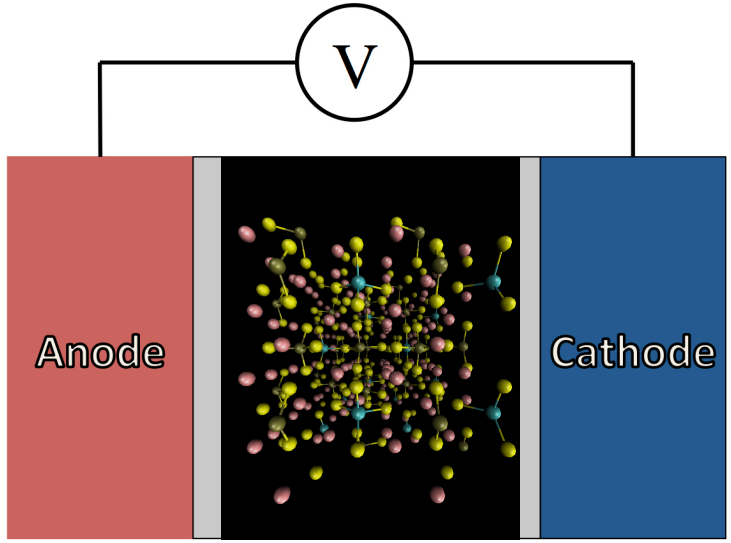
- Volatile and flammable
- Highly reactive
- Low density
- Non-dendrite suppressing

Graphite, metal

Organic solvent + lithium salt

Transition metal oxide

A major problem worth solving: liquid electrolytes



Graphite, metal

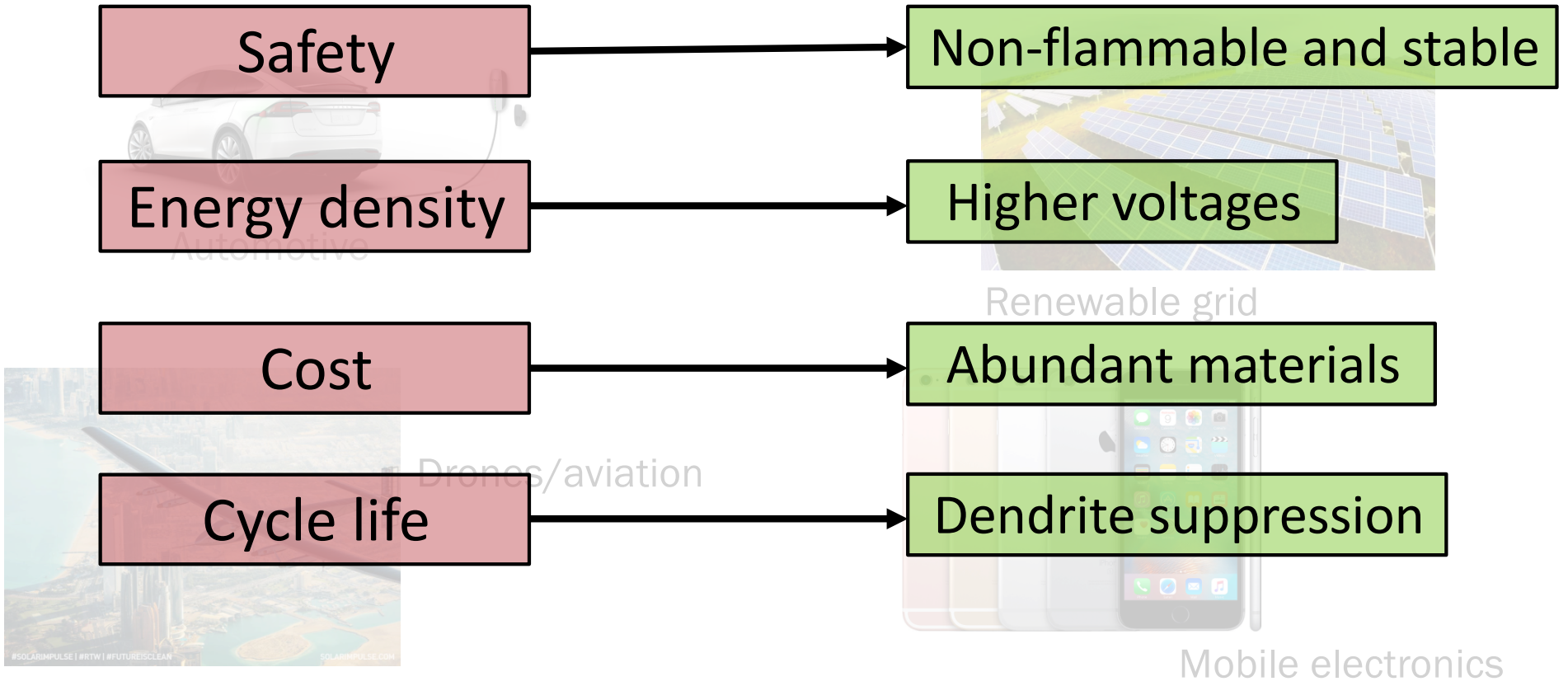
Solid crystal with mobile lithium

Transition metal oxide

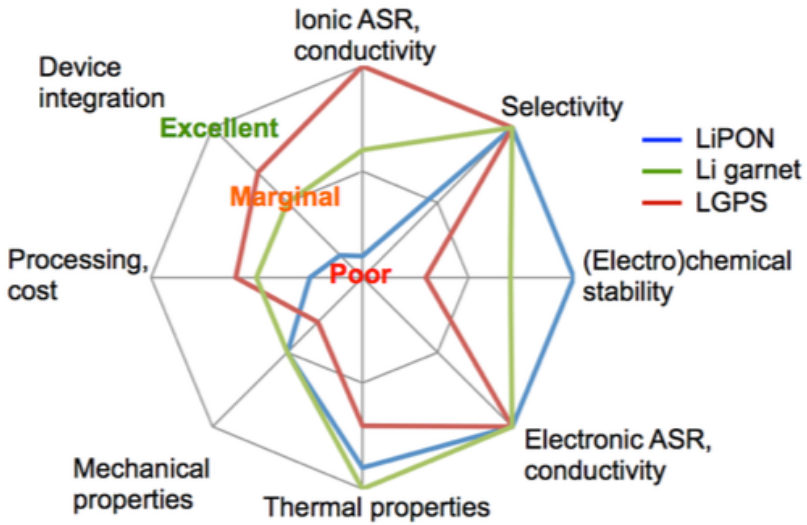
Renewable grid

Mobile electronics

Solid electrolytes represent a promising route to improvement



The challenge: trial-and-error search for candidates satisfying many criteria



Source: US Department of Energy Advanced Research Projects Agency – Energy

Electrochimica Acta, 1977, Vol. 22, pp. 773–781. Pergamon Press. Printed in Great Britain

RECENT RESULTS ON LITHIUM ION CONDUCTORS

ROBERT A. HUGGINS
Center for Materials Research, Stanford University,
Stanford, California 94305, U.S.A.

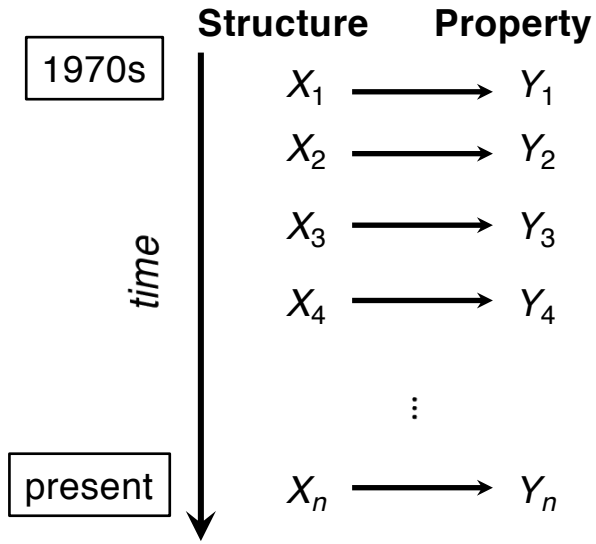
(*Received* 1 *September* 1976)

LETTERS
PUBLISHED ONLINE: 31 JULY 2011 | DOI: 10.1038/NMAT3066
nature materials

A lithium superionic conductor

Noriaki Kamaya[1], Kenji Homma[1], Yuichiro Yamakawa[1], Masaaki Hirayama[1], Ryoji Kanno[1]*, Masao Yonemura[2], Takashi Kamiyama[2], Yuki Kato[3], Shigenori Hama[3], Koji Kawamoto[3] and Akio Mitsui[4]

Existing discovery efforts are driven by trial-and-error



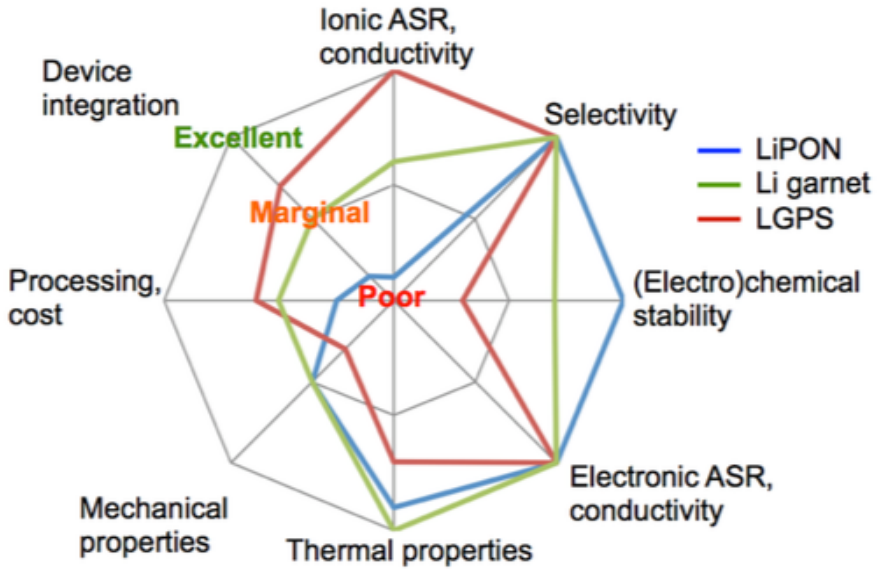
This can be formulated as a **supervised learning** problem:

$$f(X) = Y$$

f(structure) = property

Can we learn effectively from the data we've already generated?

Machine learning may improve success rate beyond trial-and-error!


Ionic ASR, conductivity
Selectivity
(Electro)chemical stability
Electronic ASR, conductivity
Thermal properties
Mechanical properties
Processing, cost
Device integration
Excellent
Marginal
Poor
LiPON
Li garnet
LGPS

We draw on the wisdom in the literature for “hypotheses”

Composition	RT bulk ionic conductivity (S cm^{-1})
$LiLa(TiO_3)_2$	1×10^{-3}
$Li_{9.81}Sn_{0.81}P_{2.19}S_{12}$	5.5×10^{-3}
$Li_{10}Ge(PS_6)_2$	1.4×10^{-2}
$Li_{10.35}Si_{1.35}P_{1.65}S_{12}$	6.5×10^{-3}
$Li_{14}ZnGe_4O_{16}(2)$	1.0×10^{-6}
$Li_2Ca(NH)_2$	6.4×10^{-6}
$Li_2Ge_7O_{15}$	5.0×10^{-6}
Li_2NH	2.5×10^{-4}
Li_2S	1.0×10^{-13}
$Li_{13.6}Si_{2.8}S_{1.2}O_{16}$	6.0×10^{-7}
$Li_{14}Ge_2V_2O_{16}$	7.0×10^{-5}
$Li_{15}Ge_3V_2O_4$	6.03×10^{-6}
$Li_{14.8}Ge_{3.4}W_{0.6}O_4$	4.0×10^{-5}
$Li_3Fe_2P_3O_{12}$	1.0×10^{-7}
Li_3N	5.75×10^{-4}
Li_3P	1.0×10^{-3}
γ-Li_3PS_4	3.0×10^{-7}
$Li_3Sc_2P_3O_{12}$	1.0×10^{-10}
β_{II}-Li_3VO_4	4.4×10^{-8}
$Li_4B_7O_{12}Cl$	1.0×10^{-7}
$Li_4BN_3H_{10}$	2.0×10^{-4}
γ-Li_4GeO_4	3.1×10^{-12}
Li_4SiO_4	2.4×10^{-10}
$Li_5La_3Bi_2O_{12}$	2.0×10^{-5}
$Li_5La_3Nb_2O_{12}$	8.0×10^{-6}
$Li_5La_3Ta_2O_{12}$	1.5×10^{-6}
Li_5NI_2	1.5×10^{-7}
$Li_6BaLa_2Ta_2O_{12}$	4.0×10^{-5}
Li_6FeCl_8	1.0×10^{-4}
Li_6NBr_3	1.5×10^{-7}
$Li_6SrLa_2Ta_2O_{12}$	7.0×10^{-6}
$Li_7La_3Zr_2O_{12}$	3.5×10^{-4}
$Li_7P_3S_{11}$	4.1×10^{-3}
$LiAlH_4$	2.0×10^{-9}
$LiAlSiO_4$	1.4×10^{-5}
$LiBH_4$	2.0×10^{-8}
LiI	1.0×10^{-6}
$LiNH_2$	4.0×10^{-10}
α'-$LiZr_2P_3O_{12}$	5.0×10^{-8}

	Feature	Pearson correlation coefficient
1	Volume per atom[a]	0.20
2	Standard deviation in Li neighbour count	0.22
3	Standard deviation in Li bond ionicity	−0.04
4	Li bond ionicity[a]	−0.18
5	Li neighbour count[a]	−0.19
6	*Li–Li bonds per Li[a]*	*0.06*
7	*Bond ionicity of sublattice[a]*	*−0.28*
8	Sublattice neighbour count[a]	−0.13
9	*Anion framework coordination[a]*	*−0.06*
10	Minimum anion–anion separation distance[a] (Å)	0.09
11	Volume per anion (Å^3)	−0.01
12	*Minimum Li–anion separation distance[a] (Å)*	*0.20*
13	*Minimum Li–Li separation distance[a] (Å)*	*−0.10*
14	Electronegativity of sublattice[a]	−0.16
15	Packing fraction of full crystal	0.16
16	Packing fraction of sublattice	0.19
17	Straight-line path width[a] (Å)	0.07
18	Straight-line path electronegativity[a]	−0.29
19	Ratio of features (4) and (7)	−0.03
20	Ratio of features (5) and (8)	−0.18
	Constant term	—

No single feature has strong correlation with ionic conductivity across the broad spectrum of 40 materials

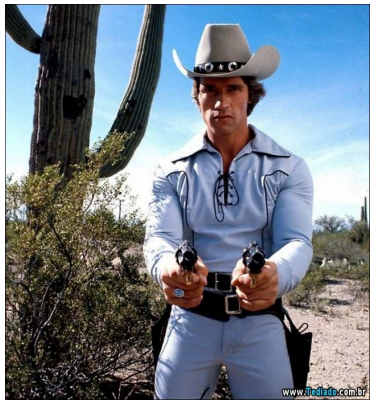
$C = corr(X, Y)$

AD Sendek, Q Yang, ED Cubuk, KAN Duerloo, Y Cui, EJ Reed. *Ener. & Environ. Sci.*, doi:10.1039/C6EE02697D (2017)

We collect 40 experimental measurements of conductivity for solids

Composition	RT bulk ionic conductivity (S cm^{-1})
$LiLa(TiO_3)_2$	1×10^{-3}
$Li_{9.81}Sn_{0.81}P_{2.19}S_{12}$	5.5×10^{-3}
$Li_{10}Ge(PS_6)_2$	1.4×10^{-2}
$Li_{10.35}Si_{1.35}P_{1.65}S_{12}$	6.5×10^{-3}
$Li_{14}ZnGe_4O_{16}(2)$	1.0×10^{-6}
$Li_2Ca(NH)_2$	6.4×10^{-6}
$Li_2Ge_7O_{15}$	5.0×10^{-6}
Li_2NH	2.5×10^{-4}
Li_2S	1.0×10^{-13}
$Li_{13.6}Si_{2.8}S_{1.2}O_{16}$	6.0×10^{-7}
$Li_{14}Ge_2V_2O_{16}$	7.0×10^{-5}
$Li_{15}Ge_3V_2O_4$	6.03×10^{-6}
$Li_{14.8}Ge_{3.4}W_{0.6}O_4$	4.0×10^{-5}
$Li_3Fe_2P_3O_{12}$	1.0×10^{-7}
Li_3N	5.75×10^{-4}
Li_3P	1.0×10^{-3}
γ-Li_3PS_4	3.0×10^{-7}
$Li_3Sc_2P_3O_{12}$	1.0×10^{-10}
β_{II}-Li_3VO_4	4.4×10^{-8}
$Li_4B_7O_{12}Cl$	1.0×10^{-7}
$Li_4BN_3H_{10}$	2.0×10^{-4}
γ-Li_4GeO_4	3.1×10^{-12}
Li_4SiO_4	2.4×10^{-10}
$Li_5La_3Bi_2O_{12}$	2.0×10^{-5}
$Li_5La_3Nb_2O_{12}$	8.0×10^{-6}
$Li_5La_3Ta_2O_{12}$	1.5×10^{-6}
Li_5NI_2	1.5×10^{-7}
$Li_6BaLa_2Ta_2O_{12}$	4.0×10^{-5}
Li_6FeCl_8	1.0×10^{-4}
Li_6NBr_3	1.5×10^{-7}
$Li_6SrLa_2Ta_2O_{12}$	7.0×10^{-6}
$Li_7La_3Zr_2O_{12}$	3.5×10^{-4}
$Li_7P_3S_{11}$	4.1×10^{-3}
$LiAlH_4$	2.0×10^{-9}
$LiAlSiO_4$	1.4×10^{-5}
$LiBH_4$	2.0×10^{-8}
LiI	1.0×10^{-6}
$LiNH_2$	4.0×10^{-10}
α'-$LiZr_2P_3O_{12}$	5.0×10^{-8}

- We adopt a binary classification strategy with a 10^{-4} S/cm boundary, motivated by engineering requirements
- Training set includes 11 “good” conductors, 29 “bad” conductors



AD Sendek, Q Yang, ED Cubuk, KAN Duerloo, Y Cui, EJ Reed. *Ener. & Environ. Sci.*, doi:10.1039/C6EE02697D (2017)

We employ logistic regression (two-class classifier)

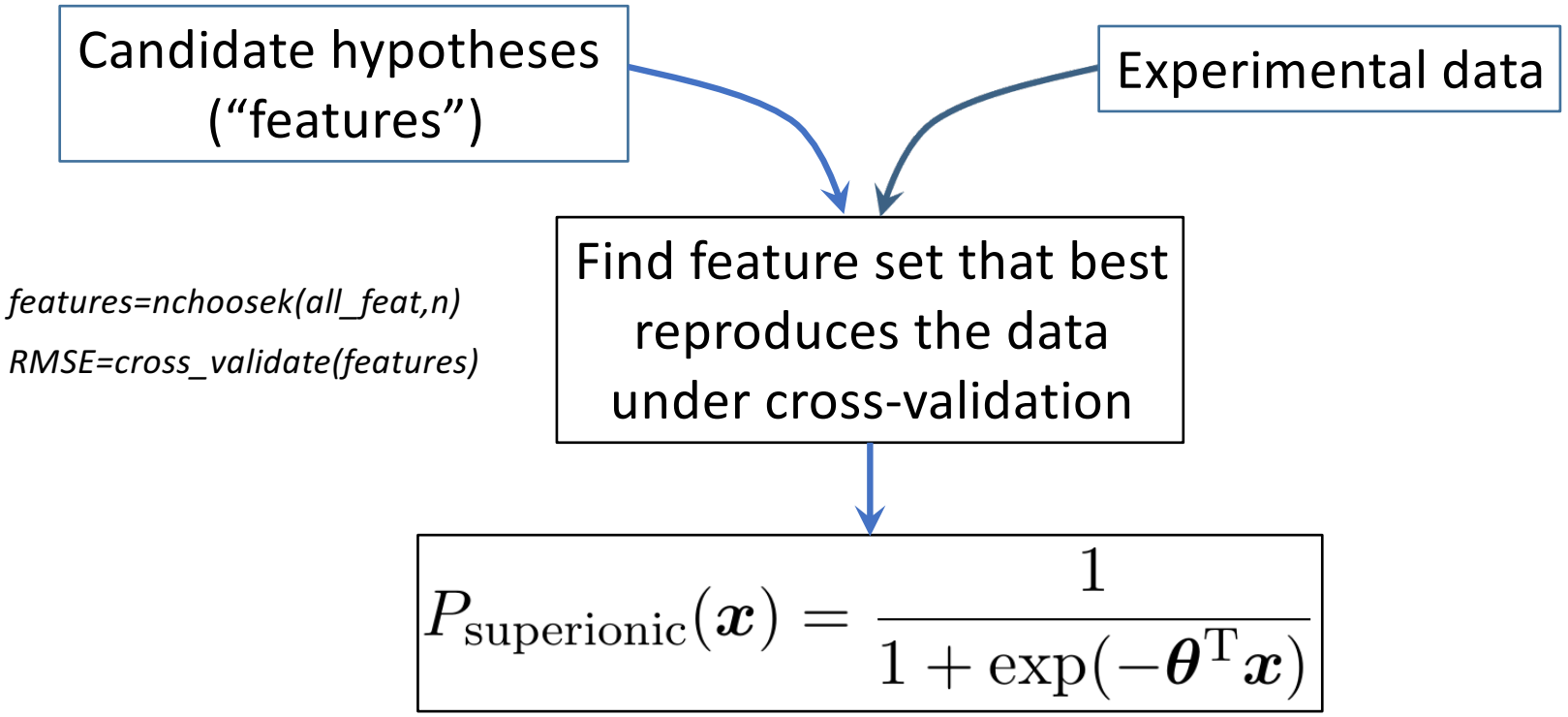
Assuming a logistic form, we search for the maximally predictive set of features:

$$P_{\text{superionic}}(\boldsymbol{x}) = \frac{1}{1 + \exp(-\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{x})}$$

$$\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{x} = ?$$

(Looping through data arrays = MATLAB's wheelhouse!)

We employ logistic regression (two-class classifier)



Machine learning finds the best Li conduction model for this data set

$$P_{\text{superionic}}(\boldsymbol{x}) = \frac{1}{1+\exp(-\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{x})}$$

$$\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{x} = 0.817LLB - 1.323SBI - 1.028AFC + 2.509LASD - 1.619LLSD - 1.944$$

Lithium-lithium bond number

Sublattice bond ionicity

Anion framework coordination

Li-anion separation distance

Li-Li separation distance

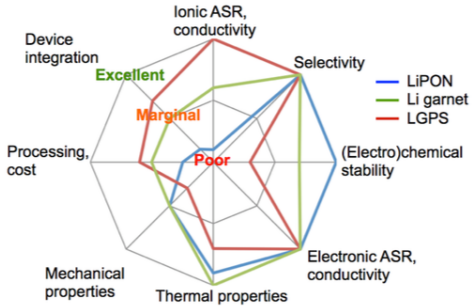
>1 week with DFT ⟶ < 1 s with ML

AD Sendek, Q Yang, ED Cubuk, KAN Duerloo, Y Cui, EJ Reed. *Ener. & Environ. Sci.*, doi:10.1039/C6EE02697D (2017)

We perform the first holistic structure screening of all >12,000 candidates

Ionic conductivity is not all that matters! We also screen on:

- High stability against oxidation ← Gibbs free energy
- High stability against reduction ← Presence of transition metals
- Low electronic conductivity ← Band gap
- High phase stability ← Convex hull
- Low cost ← Cost of raw elements involved
- High earth abundance ← Abundance of elements in Earth's crust



All Li-containing materials: >12,000

All requirements met except ionic conductivity: 317

All requirements: 21

AD Sendek, Q Yang, ED Cubuk, KAN Duerloo, Y Cui, EJ Reed. *Ener. & Environ. Sci.*, doi:10.1039/C6EE02697D (2017)

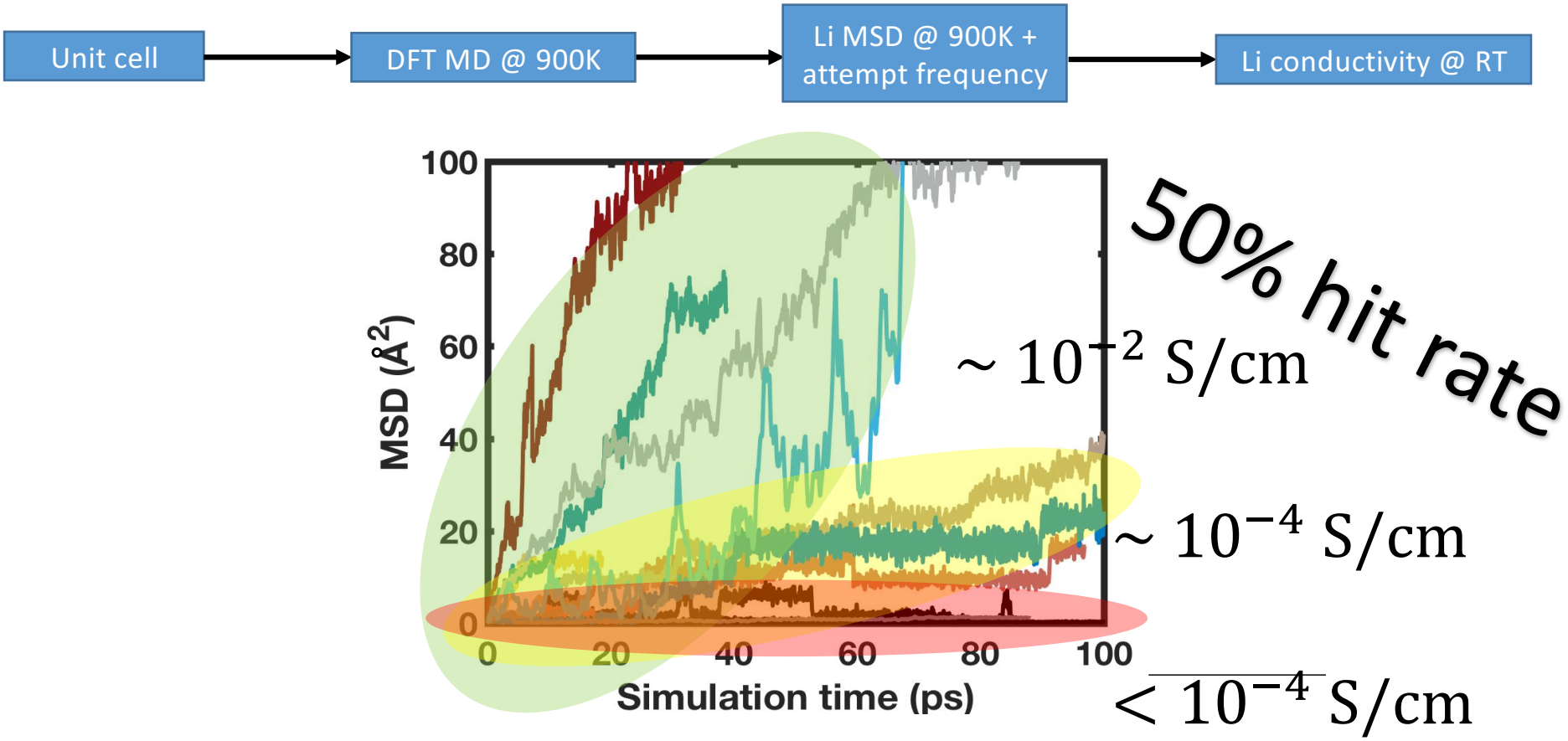
We publish 21 promising new candidates (without validation!)

MPID	Chemical formula	P_{LR}	d	ε	A	E_{gap}	$\tilde{V}_{ox}$	USD/m^2 (10 μm thick)	I_A	Related study
mp-554076	$BaLiBS_3$	0.589	1.049	0.048	1	2.153	9.697	23	38	
mp-532413	$Li_5B_7S_{13}$	0.897	1.228	0.024	1	3.553	5.454	42	38	95
mp-569782[a]	$Sr_2LiCBr_3N_2$	1.000	6.852	0.000	0	3.973	13.968	16	45	
mp-558219	$SrLi(BS_2)_3$	0.518	1.556	0.114	1	2.91	13.964	38	38	
mp-15797	$LiErSe_2$	0.543	1.505	0.056	1	1.615	6.778	170	67	
mp-29410	$Li_2B_2S_5$	0.994	1.855	0.003	1	2.538	4.895	29	38	95
mp-676361	Li_3ErCl_6	0.655	0.974	0.042	1	5.211	7.794	70	44	96 and 97
mp-643069[a]	Li_2HIO	0.652	2.081	0.079	0	4.319	4.054	2.40	60	
mp-19896	Li_2GePbS_4	0.604	1.063	0.090	1	2.265	4.591	13	54	90
mp-7744[a]	$LiSO_3F$	1.000	4.097	0.000	0	5.792	13.446	10	34	
mp-22905[b]	$LiCl$	0.837	1.381	0.031	1	6.25	4.214	0.94	34	98
mp-34477	$LiSmS_2$	0.89	1.33	0.028	1	1.921	8.536	6.50	40	
mp-676109	Li_3InCl_6	0.656	1.013	0.058	1	3.373	6.215	5.50	63	96 and 97
mp-559238	$CsLi_2BS_3$	0.812	1.642	0.055	1	3.094	4.798	160	49	
mp-866665[a]	$LiMgB_3(H_9N)_2$	1.000	5.149	0.000	0	6.511	11.222	30	38	
mp-8751	$RbLiS$	0.775	1.279	0.051	1	2.745	4.22	240	34	
mp-15789	$LiDyS_2$	0.901	1.339	0.025	1	1.935	8.736	9.20	39	
mp-15790	$LiHoS_2$	0.899	1.327	0.025	1	1.965	8.749	300	55	
mp-15791	$LiErS_2$	0.899	1.319	0.025	1	2.008	8.761	190	44	
mp-561095[a]	$LiHo_3Ge_2(O_4F)_2$	0.984	3.247	0.009	0	4.163	53.18	370	55	
mp-8430	$KLiS$	0.76	1.243	0.052	1	3.057	4.348	14	34	

…and hope we're right

How well does the model do?

We discover ten new solids that are superionic conductors



AD Sendek, ED Cubuk, Y Cui, EJ Reed. In preparation. (2017)

We discover ten new solids that are superionic conductors

ML-guided simulations

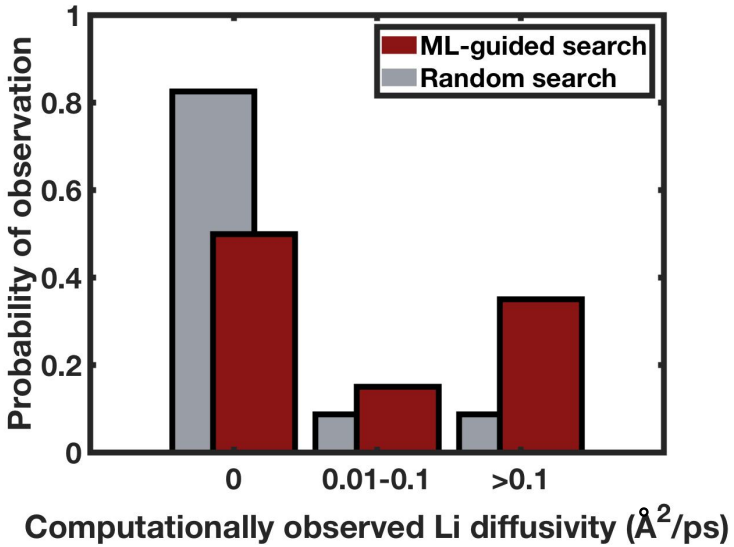
Name	$P_{superionic}$
$BaLiBS_3$	58.9%
$Li_5B_7S_{13}$	89.7%
$Sr_2LiCBr_3N_2$	100%
$SrLi(BS_2)_3$	51.8%
$LiErSe_2$	54.3%
$Li_2B_2S_5$	99.4%
Li_3ErCl_6	65.5%
Li_2HIO	65.2%
Li_2GePbS_4	60.4%
$LiSO_3F$	100%
$LiCl$	83.7%
$LiSmS_2$	89.0%
Li_3InCl_6	65.6%
$CsLi_2BS_3$	81.2%
$LiMgB_3(H_9N)_2$	100%
$RbLiS$	77.5%
$LiDyS_2$	90.1%
$LiHoS_2$	89.9%
$LiErS_2$	89.9%
$LiHo_3Ge_2(O_4F)_2$	98.4%
$KLiS$	76%

Model success rate: ~35%
(melting cases excluded)

Random success rate: ~10%
(melting cases excluded)

Random simulations

Name	$P_{superionic}$
$LiLa_2SbO_6$	0
Li_6UO_6	7.70%
$Li_4H_3BrO_3$	43%
$LiInF_4$	0.60%
$LiBiF_4$	0.20%
$CsLi_2(HO)_3$	3.50%
$Li_6Ho(BO_3)_3$	10.10%
$RbLiB_4O_7$	4%
$LiU_4P_3O_{20}$	0.10%
$Li_4Be_3As_3ClO_{12}$	0.30%
Li_6TeO_6	6.40%

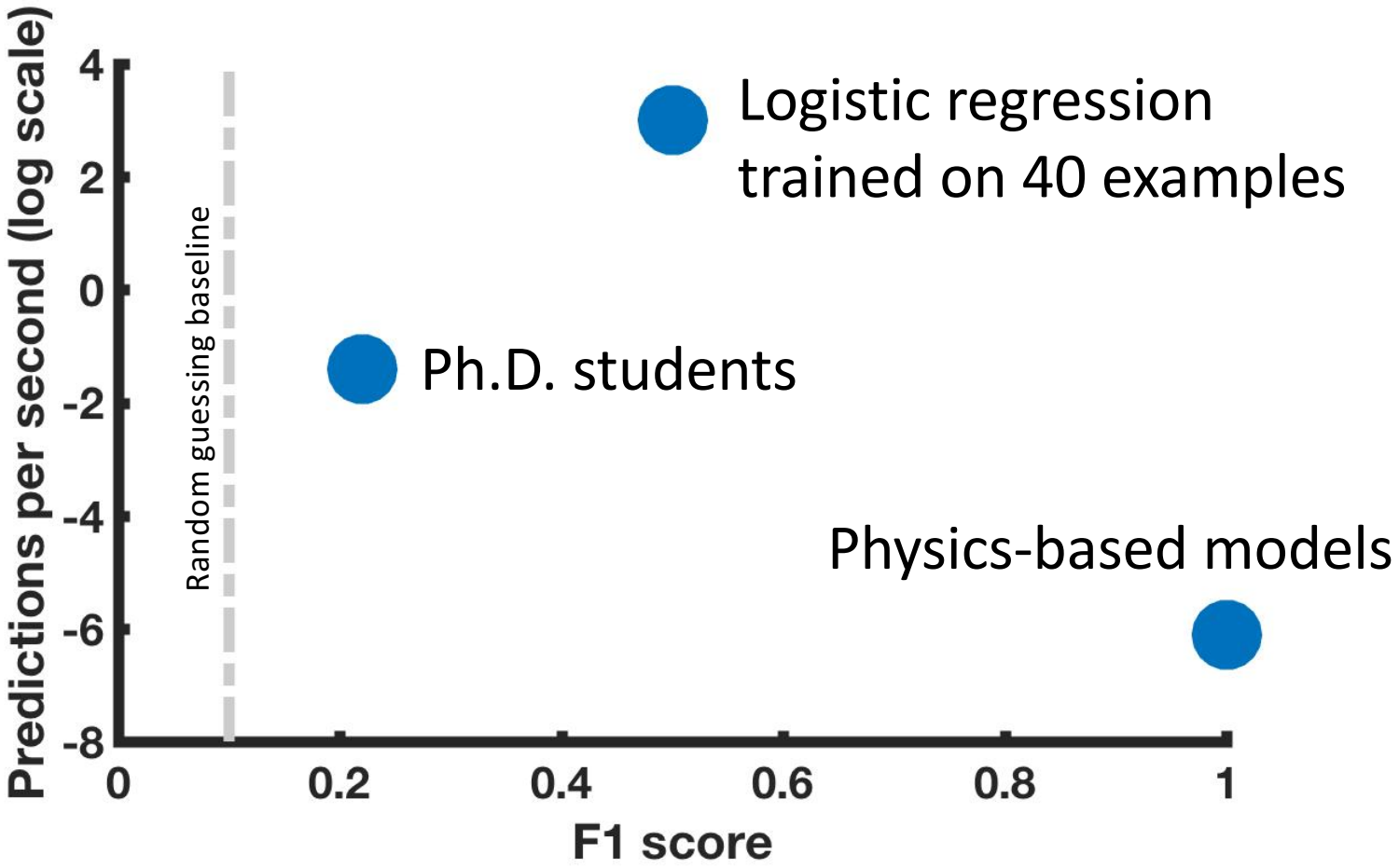


Learning from only 40 data points = 3x improvement over guesswork

+ exciting, new, potentially record-setting material!

AD Sendek, ED Cubuk, Y Cui, EJ Reed. In preparation. (2017)

Our algorithm outperforms humans in accuracy and speed



AD Sendek, ED Cubuk, G Cheon, Y Cui, EJ Reed. In preparation. (2017)

We leverage machine learning to identify the most promising solid electrolytes 10^6x faster than state-of-the-art

- Our model is ~3x better than random guessing
- Achieved by collecting most of the available data and proposed wisdom over past four decades
- We perform the first holistic screening of all known Li containing solids and discover several new superionic conducting structures
- Exciting new materials for further development!

Screening: AD Sendek, Q Yang, ED Cubuk, KAN Duerloo, Y Cui, EJ Reed. *Ener. & Environ. Sci.*, doi:10.1039/C6EE02697D (2017).
DFT MD: AD Sendek, ED Cubuk, Y Cui, EJ Reed. In preparation.

40 data points = 3x improvement

Imagine what we can do with 100, 500, 1000?

The value we get from MATLAB

Quick prototyping

- MATLAB is easy and intuitive to use; get a prototype out the door faster

Text file processing

- Very easy to scan through directories and read in files

Thorough, up-to-date documentation

- Can always get answers quickly

Figure generation

- All our figures are made in MATLAB; it is easy and the plots are clean

TomKatCenter
FOR SUSTAINABLE ENERGY
STANFORD UNIVERSITY

Thank you!

Austin Sendek
Email: *asendek@stanford.edu*
Web: *stanford.edu/~asendek*

Prof. Evan Reed: evanreed@stanford.edu