MAGICIAN 養成講座

MAterials Genome/Informatics and Chemo-Informatics Activate Networks

第 5 回データのクレンジング

—Li イオンの拡散係数を例に—

2018.8.28　横浜国大非常勤講師　山本博志

マテリアル・ジーノム、マテリアル・インフォマティクスなどの世界で、収集されたデータのクレンジングなどという言葉が出てくる。最初は意味が解らなかった。家庭内にはクレンザー（汚れ落とし）があるし、お化粧を落とす事をクレンジングなどとも言うが、このクレンジングは和製英語で、英語には化粧を落とすと言う意味は無い。英語の Cleansing（クレンジング）は洗浄すると言う意味だ。データをクレンジングすると言った場合、英語の本来の意味ではなく、DB 中の汚いデータを取り除くと言う和製英語の意味に使っているように思える。毒性などの特殊な DB を除き、物性値などの DB は、どういう研究機関の誰が、何時、どういう方法で測定したか、論文などのソースなどが記載されている。同じ対象であっても様々に異なるデータが記載されていることもある。古いデータなどでは単位が異なっていたり、測定条件があやふやだったりする。データを”クレンジング”してからでなければ、そのままマテリアル・ジーノムで解析というわけにはいかない。ネット上から Li イオンの拡散のデータが得られたので、これを例題に、データの前処理の方法を説明しておこう。

Li イオンの拡散データ

たまたまネットを調べていたところ、

“リチウム複合酸化物におけるイオン伝導度の系統的評価”京都大学、炭谷晃史 博士論文　2017-03-23

Dkogk04307.pdf

という論文が見つかった。リチウム複合酸化物 628 物質のリチウムイオン拡散係数の値が記載されている。原本のテーブルは容易に表計算ソフトにインポートすることができる。複合酸化物の、各原子の数はとりあえず手入力で列に入れる（所要時間 1 時間）。

B	C	D	E	F	G	H	AO	AP	AQ	AR
Chemical formula	Log D_{Li}	Li	Al	Si	O	Nb	Tl	Cd	Tb	Sum
$LiAlSiO_4$	-7.665	1	1	1	4	0	0	0	0	7
$LiAlSiO_4$	-7.777	1	1	1	4	0	0	0	0	7
$LiAlSiO_4$	-7.902	1	1	1	4	0	0	0	0	7
$Li_{0.9}AlSiO_4$	-7.993	0.9	1	1	4	0	0	0	0	6.9
$LiAlSiO_4$	-8.045	1	1	1	4	0	0	0	0	7
$Li_5Nb_2La_3O_{12}$	-8.086	5	0	0	12	2	0	0	0	22
Li_2TiO_3	-8.102	2	0	0	3	0	0	0	0	6

次に組成をパーセント表記にする。

AS	AT	AU	AV	AW
Li	Al	Si	O	Nb
14.3	14.3	14.3	57.1	0.0
14.3	14.3	14.3	57.1	0.0
14.3	14.3	14.3	57.1	0.0
13.0	14.5	14.5	58.0	0.0
14.3	14.3	14.3	57.1	0.0
22.7	0.0	0.0	54.5	9.1

これは、複合酸化物の表記の違いを吸収するために必要だ。つまり、$LiAlSiO_4$ と $Li_2Al_2Si_2O_8$ は同じものにするためだ。

論文中の明らかな誤記、打ち込みミス

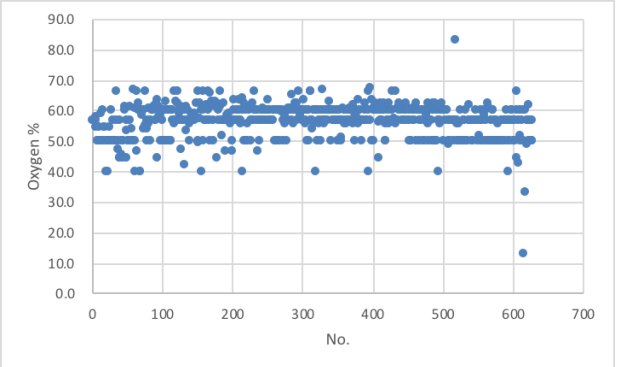
$Li_{0.3335}Fe_{0.3335}Zn_{0.3335}O_5$

$Li_{0.05}Fe_{25}Zn_{0.9}O_4$

O_5 や Fe_{25} はあり得ないのですぐわかるが、容易に修正可能でなければ、とりあえず除外しておく。

どうすれば、こうした間違いを発見できるだろうか?

地道に、出来上がったテーブルの列を地道にプロットしてみる。



酸素の含有量が 33%以下、67%以上は間違いが無いかチェックする事が必要だろう。酸素が一番少ない状況は、Li_2O などで 33%, 一番多い状況は、SiO_2 などで 66.6%である事を理解している必要がある。

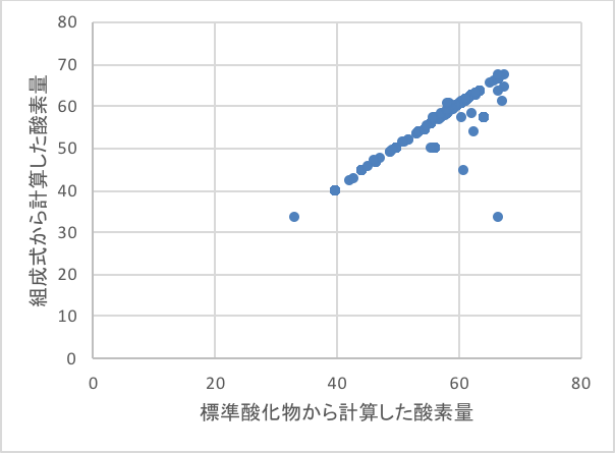
複合酸化物の組成式から原子の列を作る際に生じた入力ミスも、こうしたプロットをしてみると発見される事が多

い。

次に、標準酸化物を検討する。

原子のうち遷移金属は酸化数の異なる酸化物が存在する。その中で、最も一般的な酸化数の酸化物で複合酸化物を作った場合の理想的な酸素含有量を計算する。

原子	標準酸化物	酸素係数			
Li	Li2O	0.5			
Al	Al2O3	1.5			
Si	SiO2	2			
Nb	Nb2O5	2.5	Nb2O5	NbO	NbO2
La	La2O3	1.5			
Ti	TiO2	2	TiO2	TiO	Ti2O3
Sn	SnO2	2	SnO2	SnO	
Zr	ZrO2	2	ZrO2		
Ce	CeO2	2	CeO2	Ce2O3	
Hf	HfO2	2	HfO2		
P	P2O5	2.5			
Fe	FeO	1	Fe2O3	Fe3O4	
K	K2O	0.5			
Ge	GeO2	2	GeO		
Mn	MnO2	2	MnO	Mn2O3	Mn2O7
V	VO2	2	V2O3	VO2	V2O5
Zn	ZnO	1			
Ca	CaO	1			
In	In2O3	1.5			
Sr	SrO	1			
Gd	Gd2O3	1.5			
Rb	Rb2O	0.5			
Ni	NiO	1	Ni2O3		
Ga	Ga2O3	1.5	Ga2O		
Mg	MgO	1			
Eu	Eu2O3	1.5			
Co	CoO	1	CoO	Co3O4	Co2O3
Er	Er2O3	1.5			
Ho	Ho2O3	1.5			
Y	Y2O3	1.5			
B	B2O3	1.5			
Yb	Yb2O3	1.5			
Nd	Nd2O3	1.5			
Na	Na2O	0.5			
Ba	BaO	1			
Cu	Cu2O	0.5	CuO		
Tl	Tl2O2	2	Tl2O3	Tl2O	
Cd	CdO	1			
Tb	Tb2O3	1.5	Tb4O7		



本来は、酸素量は一致しなくてはならないが、主に打ち込みミスによって、値が異なるので、テーブルを修正する。修正が終わっても上図に示すように直線から大きく乖離するデータがある。これらはほとんどFe化合物である。鉄にはFeO, Fe_2O_3, Fe_3O_4の形があり、今回はFeOを「標準形」と置いたが、Fe_2O_3を原料に用いている場合には、当然、酸化数が異なった複合酸化物になる。それ以外、微量の酸素欠陥、酸素過剰の複合酸化物も存在することがわかる。これは、複合酸化物を作る際に、仕込み組成で表示するか、できた複合酸化物の組成を機器分析した結果を表示するかで微妙に変わる可能性がある。酸化物によっては、融点、沸点が低く調整中に揮散する事もあるし、調整する雰囲気によっても変化する可能性がある。

ここまでクレンジングが終わると、予測式の検討に入ることができる程度のデータセットとなる。

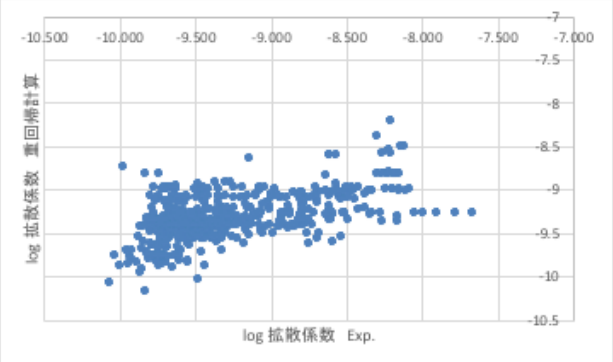
予測式の（予備）検討

作成したテーブルを元に、まず、重回帰計算を行う。

log 拡散係数 = 0.015380859375*Li +0.009765625*Al -0.002197265625*Si +0.00390625*O+……

という予測式が求まる。

この重回帰式を求める機能はExcelにも搭載されているのでぜひ試して見ていただきたい。重回帰式は非常に見通しの良い式で、例えば、Siが多くなると、係数がマイナスなので拡散係数は小さくなるとか、Liは係数が非常に大きいだとかがわかる。ただし、重回帰法は、項目間に相互作用がある場合、現象に非線形性がある場合には予測性能はとても劣るという特徴を持っている。



今回のデータ全てをプロットすると上図のようになる。これでは、ある複合酸化物があった時に、その予測値を得るのは不可能であると言わざるを得ない。

それなのに、なぜ敢えて重回帰計算をするのかというと、もう一つのクレンジングを行うためである。重回帰計算値

でソートをかけてみよう。すると、計算値が全く同じになるものが存在することがわかる。例えば LiAlSiO4 という複合酸化物は同じ組成ながら、拡散係数の実験値が-7.665 から-9.621 と 100 倍異なるデータが存在することがわかる。計算値は組成が同じなので同じ値-9.261 になる。したがって推算式は精度が出なくて当たり前になる。

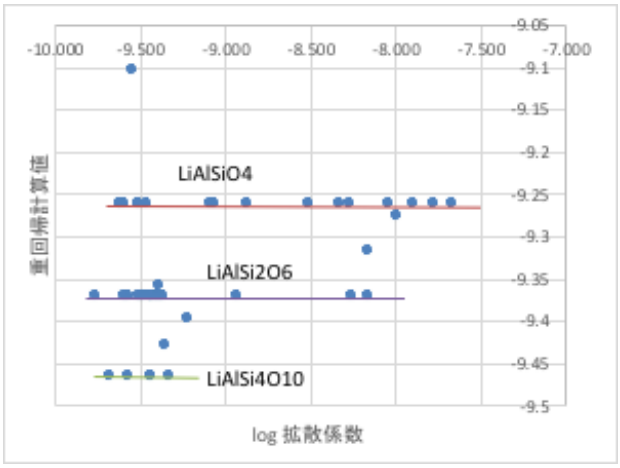
ICSD No.	Chemical formula	Log D_{Li}
42499	$LiAlSiO_4$	-7.665
32595	$LiAlSiO_4$	-7.777
22015	$LiAlSiO_4$	-7.902
2929	$LiAlSiO_4$	-8.045
22011	$LiAlSiO_4$	-8.269
22010	$LiAlSiO_4$	-8.334
22014	$LiAlSiO_4$	-8.510
66137	$LiAlSiO_4$	-8.870
97910	$LiAlSiO_4$	-9.509
30982	$LiAlSiO_4$	-9.511
67237	$LiAlSiO_4$	-9.594
30983	$LiAlSiO_4$	-9.619
92708	$LiAlSiO_4$	-9.621

それでは、重回帰計算の結果の-9.261 付近のデータを残して、他の値をクレンジングしてしまえば良いのかというとそう単純な話では無い。log の拡散係数が-7.665 というのは、このテーブルの中で最も拡散係数が高いものである。それを重回帰計算に合わないからと言って、消去してしまうと本来欲しい高い拡散係数の組成が失われることになる。

テーブルの中には、Li-Al-Si-O 系としては、次の 10 種類の複合酸化物が存在して、データ数は 52 件存在する。

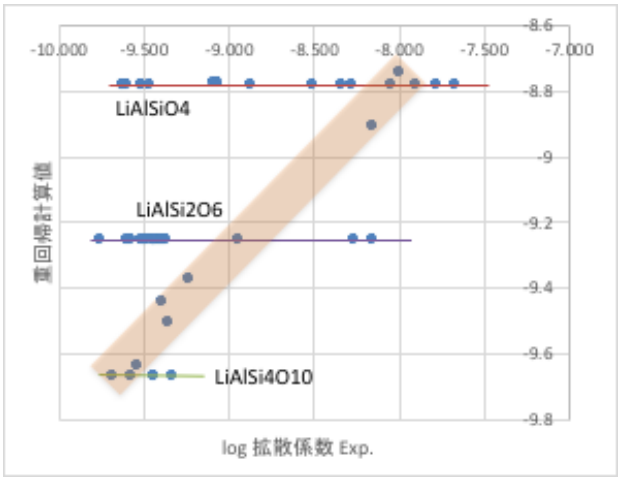
Chemical formula
$Li_{0.9}AlSiO_4$
$Li_{1.146}Al_{1.32}Si_{1.68}O_6$
$Li_{2.5}Al_{0.5}SiO_4$
$Li_{3.53}Al_{3.53}Si_{8.47}O_{24}$
$Li_{95.68}Al_{96}Si_{96}O_{384}$
$Li_{96}Al_{96}Si_{96}O_{384}$
$LiAlSi_2O_6$
$LiAlSi_3O_8$
$LiAlSi_4O_{10}$
$LiAlSiO_4$

先ほど作成した、重回帰計算のうち、Li-Al-Si-O 系だけを取り出すと下図のようになる。



組成が同じであれば、計算値が同じになるが、その周辺の組成の値と大きく異なる値を残すのは合理的では無い。

新たに、Li-Al-Si-O 系だけで重回帰計算を行い、データのクレンジングを行う。



残すデータは、オレンジ色の領域のデータとするのが合理的であると言えるだろう。

こうした、どのデータを残し、どのデータを削除するのかは、研究者によって差が出る所だ。また、Li-Al-Si-O 系のように、いつも合理的にクレンジングできるわけでも無い。場合によると、クレンジングせずにデータを残して置いて、平均値を使うというという選択肢も大事なセンスになる。

Li-Ti-La-O 系では多くの系をクレンジングして消去したが、どうしても判断のつかない系は残した。

ICSD No.	Log D_{Li}	Type	O	Li	La	Ti	Chemical formula
59647	-9.185	LiTiLaO	62.7615063	3.34728033	12.9707113	20.9205021	$Li_{0.16}TiLa_{0.62}O_3$
82670	-9.590	LiTiLaO	62.4089869	4.36862908	12.4193884	20.8029956	$Li_{0.21}TiLa_{0.597}O_3$
172044	-9.240	LiTiLaO	61.6396137	6.16396137	11.649887	20.5465379	$Li_{0.3}TiLa_{0.567}O_3$
163219	-9.169	LiTiLaO	63.1578947	2.63157895	13.1578947	21.0526316	$Li_{0.125}TiLa_{0.625}O_3$
163220	-8.898	LiTiLaO	63.1578947	2.63157895	13.1578947	21.0526316	$Li_{0.125}TiLa_{0.625}O_3$
155634	-8.783	LiTiLaO	61.2244898	7.14285714	11.2244898	20.4081633	$Li_{0.35}TiLa_{0.55}O_3$
155633	-8.738	LiTiLaO	61.2244898	7.14285714	11.2244898	20.4081633	$Li_{0.35}TiLa_{0.55}O_3$
165469	-9.303	LiTiLaO	60	10	10	20	$Li_{0.5}TiLa_{0.5}O_3$
82907	-9.441	LiTiLaO	58.8235294	11.7647059	11.7647059	17.6470588	$Li_2Ti_3La_2O_{10}$
81857	-9.622	LiTiLaO	57.1428571	14.2857143	14.2857143	14.2857143	$LiTiLaO_4$
411039	-9.414	LiTiLaO	60.1202405	9.73947896	10.1002004	20.0400802	$Li_{0.486}TiLa_{0.504}O_3$
50434	-9.431	LiTiLaO	60.6060606	8.68686869	10.5050505	20.2020202	$Li_{0.86}Ti_2La_{1.04}O_6$
50435	-9.196	LiTiLaO	61.6016427	6.36550308	11.4989733	20.5338809	$Li_{0.62}Ti_2La_{1.12}O_6$
50436	-8.805	LiTiLaO	63.1578947	2.73684211	13.0526316	21.0526316	$Li_{0.26}Ti_2La_{1.24}O_6$

組成が整数比でないものは、結晶系などの違いによって特異的に結果が変わりうるので、無視はできない事もある。

特に、これらの Li-Ti-La-O 系では顕著になるのだが、標準酸化物で計算した酸素量と、実際の酸素量が一致しない。原子の組成比に加えて、酸素の過剰、欠乏も学習データに加えることにする。

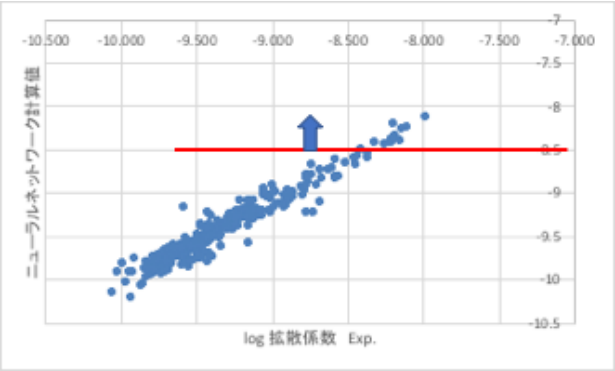
今回は、複合酸化物として 628 系のデータがあったが、最終的には 251 系と半分以下のデータを用いて最終的な予測式を作成した。このようにビッグデータを使うよりも、質の高いデータを選び出す事の方が成功に繋がるように思える。

予測式の作成

ここではニューラルネットワーク(NN)法を用いて推算式を構築する。重回帰法と比べ、NN 法は項目間の相互作用が取り込めたり、非線形性を加味された予測式を作るには適している。ただし、できあがった予測式の情報の流れがブラックボックス化してしまい化学的な解析がしにくいのが欠点でもある。NN 法による物性推算に関しては別の章で詳しく解説しようと思う。

NN 法のソフトウエアーに関しては、最近は機械学習のソフトウエアーとして様々なものが流通しているので、適当なものを探して使って欲しい。学生用のバージョンは所定のサイトからダウンロードして利用する。

データのクレンジングがうまく行っていれば、比較的簡単に収束して、組成を入れれば予測値を返すシステムが構築できる。多くの NN 法では、中間層のニューロン数の設計が一番難しい点となる。ネットワークの構造を変えただけでも結果は大きく変化する。



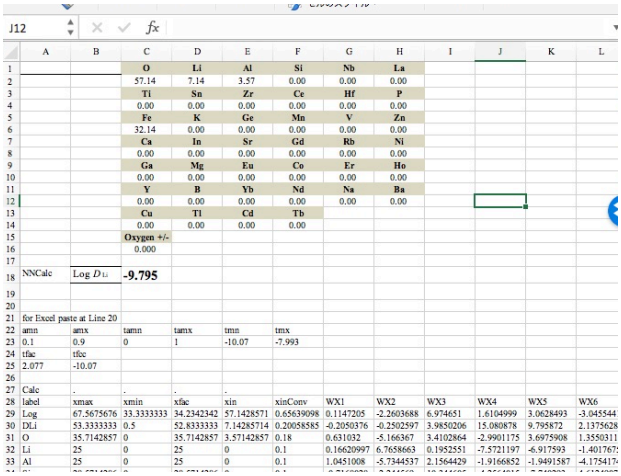
最終的に採用した NN は上図のようなパーフォーマンスを持つものだ。相関係数的にはさらに高いものも作れるが、その分、非線形性が高まり、予測性能が逆に悪くなる。なれるまでは中間層のニューロンをなるべく減らしたほうが良い結果が得られる。

NN 法で予測値が-8.5 より大きくなる（赤線よりも上）構成を押さえておく。

4 元系		3 元系
LiAlSiO		LiSnO
LiZrLaO		LiCeO
LiNbLaO		LiTiO
LiLaHfO		LiAlO
LiSnLaO		
5/85系		4/22系

4 元系で 5 種類、3 元系で 4 種類の複合酸化物があることがわかる。

授業で使うニューラルネットワーク法のソフトウエアーはエクセルの表形式でも出力されるので，複合酸化物の組成を入力すれば自動的に予測値を答える計算機を Excel 上に構築することができる。酸素の過剰、欠乏は自動的に計算するので入力する必要はない。



	A	B	C	D	E	F	G	H
1			O	Li	Al	Si	Nb	La
2			57.14	7.14	3.57	0.00	0.00	0.00
3			Ti	Sn	Zr	Ce	Hf	P
4			0.00	0.00	0.00	0.00	0.00	0.00
5			Fe	K	Ge	Mn	V	Zn
6			32.14	0.00	0.00	0.00	0.00	0.00
7			Ca	In	Sr	Gd	Rb	Ni
8			0.00	0.00	0.00	0.00	0.00	0.00
9			Ga	Mg	Eu	Co	Er	Ho
10			0.00	0.00	0.00	0.00	0.00	0.00
11			Y	B	Yb	Nd	Na	Ba
12			0.00	0.00	0.00	0.00	0.00	0.00
13			Cu	Tl	Cd	Tb		
14			0.00	0.00	0.00	0.00		
15			Oxygen +/-					
16			0.000					
17								
18	NNCalc	Log D_{Li}	-9.795					
19								
20								
21	for Excel paste at Line 20							
22	amn	amx	tamn	tamx	tmn	tmx		
23	0.1	0.9	0	1	-10.07	-7.993		
24	tfac	tfec						
25	2.077	-10.07						
26								
27	Calc	.	.	.	.			
28	label	xmax	xmin	xfac	xin	xinConv	WX1	WX2
29	Log	67.5675676	33.3333333	34.2342342	57.1428571	0.65639098	0.1147205	-2.2603688
30	DLi	53.3333333	0.5	52.8333333	7.14285714	0.20058585	-0.2050376	-0.2502597
31	O	35.7142857	0	35.7142857	3.57142857	0.18	0.631032	-5.166367
32	Li	25	0	25	0	0.1	0.16620997	6.7658663
33	Al	25	0	25	0	0.1	1.0451008	-5.7344537

	I	J	K	L
1				
2				
3				
4				
5				
6				
7				
8				
9				
10				
11				
12				
13				
14				
15				
16				
17				
18				
19				
20				
21				
22				
23				
24				
25				
26				
27				
28	WX3	WX4	WX5	WX6
29	6.974651	1.6104999	3.0628493	-3.045544
30	3.9850206	15.080878	9.795872	2.1375628
31	3.4102864	-2.9901175	3.6975908	1.3550311
32	0.1952551	-7.5721197	-6.917593	-1.401767
33	2.1564429	-1.9166852	-1.9491587	-4.175417

Excel でマクロを書くことができれば，例えば Li の量を変化させた時に，拡散係数がどうなるか，連続計算させることも容易である。

O	Li	Al	Si	Nb		Log D_{Li}Calc
57.14	**14.29**	14.29	14.29	0		**-8.972**
57.14	**14.24**	14.29	14.29	0		**-8.942**
57.14	**14.19**	14.29	14.29	0		**-8.912**
57.14	**14.14**	14.29	14.29	0		**-8.881**
57.14	**14.09**	14.29	14.29	0		**-8.849**
57.14	**14.04**	14.29	14.29	0		**-8.817**
57.14	**13.99**	14.29	14.29	0		**-8.784**

ただし，基本形から大きく乖離する組成を計算することはお勧めしない。これらの複合酸化物は結晶化ガラスと呼ばれるガラスである。基本形から大きく外れるものは結晶の構造が大きく変化し予測する拡散係数が大きく外れることもある。

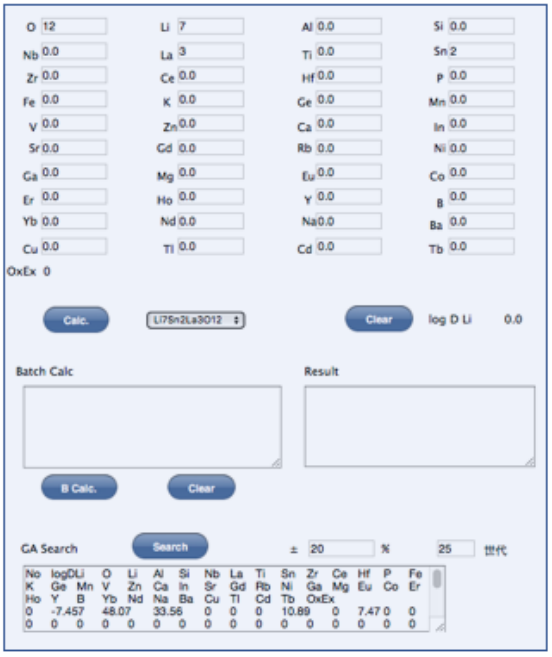
また，予測式は JAVA のプログラムとしても出力されるので，プログラムが書ける場合には，より高度なアプリケーションを作成することも可能である。

こうした計算機で，より高い拡散係数となると予測される新たな組成を実験し，結果を予測式にフィードバックし，さらに高い拡散係数となる組成を求めて行く。予測式は一度作ったら終わりではなく，日々更新して行くものである。Excel や Java の必要な部分だけを差し替え、バージョンを上げていかれる手軽さは非常に重要である。

逆設計

遺伝的アルゴリズムを用いた逆設計については，別の章で詳しく説明した。ここでは，次の機能を持つ JavaScript のプログラムを作成した。

1. 拡散係数を予測する。
2. 連続計算を行う。
3. 遺伝的アルゴリズムを使って，基本形の組成を少し動かして，性能を向上させる組成設計を行う



このプログラムは，ブラウザー上で動作するので，PC だけではなく，スマートフォン上でも動作する。

基本組成を入力して，組成を動かす%を入力して探索を行うと改良処方が提示される。基本処方は、手入力も可能だが、NN 法で予測値が-8.5 より大きくなるものはプルダウンメニューからも選択が可能になっている。
こうして得られた処方を元に、ある原子だけ量を増減して連続計算するルーチンも搭載してある。

No		logDLi	O	Li	Nb	La	Ti	Sn	Zr	Hf
Li7Sn2La3O12	実験値	-8.208	50.00	29.17		12.50		8.33		
GA探索	予測値	-7.359	47.64	34.25		11.38		6.72		
No		logDLi	O	Li	Nb	La	Ti	Sn	Zr	Hf
Li7La3Hf2O12	実験値	-8.221	50.00	29.17		12.50				8.33
GA探索	予測値	-7.667	47.56	34.43		11.33				6.67
No		logDLi	O	Li	Nb	La	Ti	Sn	Zr	Hf
Li7Zr2La3O12	実験値	-8.148	50.00	29.17		12.50			8.33	
GA探索	予測値	-7.749	45.03	40.74		7.61			6.62	
No		logDLi	O	Li	Nb	La	Ti	Sn	Zr	Hf
Li5Nb2La3O12	実験値	-8.086	54.55	22.73	9.09	13.64				
GA探索	予測値	-7.588	50.71	32.75	9.52	7.02				
No		logDLi	O	Li	Nb	La	Ti	Sn	Zr	Hf
Li2SnO3	実験値	-8.139	50.00	33.33				16.67		
GA探索	予測値	-7.778	48.41	36.51				15.08		

予測式の拡張

今回のデータソースには３元系，４元系の複合酸化物し

か含まれていない。ニューラルネットワーク法などで物性を予測する場合，基本的には外挿性は期待してはいけない。つまり今回作成したプログラムで，5,6 元系の複合酸化物を計算してはいけない。

それでは計算は常に実験の後追いになってしまうので面白みに欠けることになる。

そうした時には，別のソースを当たってみることも重要である。例えば InterGlad というガラスのデータベースがある。そこには，様々なガラスの組成と物性値が記載されている。その中で電気伝導度のデータを調べ，電気伝導度の高い組成を抜き出してくる。

Si-Li-O
Li-Mo-O
B-Li-O
Si-Al-Ca-Li-Ti-Ni-O
B-Li-Na-O
B-Li-K-O
Al-Li-Ti-P-O
Li-Ti-P-Ga-O

こうした組成を中心に予測式の拡張を図るのも一つの手である。

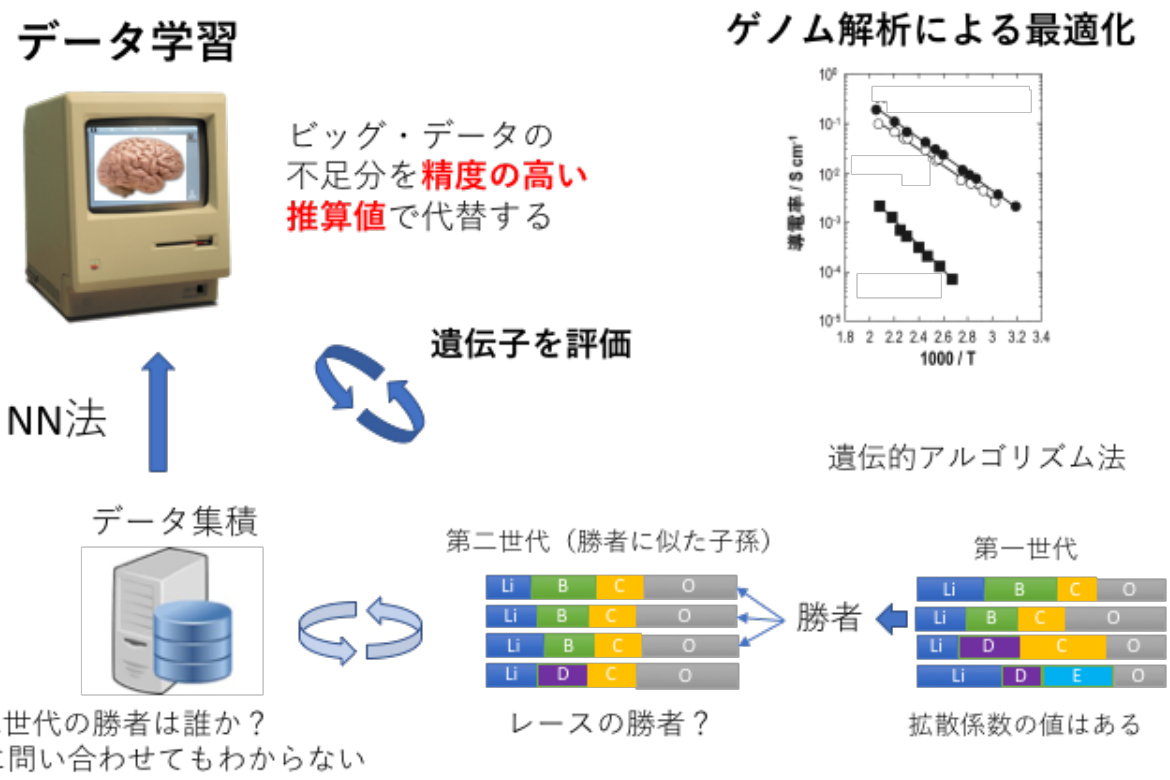
このデータベースには，

Li_2S-GeS_2-LiI

Li_2S-LiI-SiS_2

Li_2S-P_2S_5

などの硫化物系の組成も記載されているので，さらに広い範囲に拡張が可能であろう。このデータベースを使うのであれば，再び，クレンジングから地道に繰り返す必要がある。そして，拡散係数と電気伝導度の両面から組成設計を行って行くのがマテリアル・ジーノムの一番面白いところと言える。



雑感

マテリアル・ジーノムを使うと実験などしなくても新しい素材が次々と生まれるような誤解があるように思える。また、スーパーコンピュータや高価なソフト、高額な分析装置が必要であるという誤解があるように思える。そしてそれを使う側にも、高度な数学、統計解析能力が必要であ

るという誤解があるように思える。その敷居の高さが普及を遅らせているようだ。ニューラルネットワークに化学を教えようとした時に、一番適した教育者は化学者自身である。本来、化学の知識を元にクレンジングすれば少ないデータ数でも精度は上がるはずである。それを、中をブラックボックス化して、大量のデータを用いて行うのが現状のマテリアル・ジーノムとなっている。データサイエンティストに化学を学ばせるのが早いか、化学者がちょっとしたテクニックを覚えて少ないデータで効率的なマテリアル・ジーノムを実行するか。どちらが良いかは明らかなように思える。その上で、問題に突き当たったら、赤池情報量基準を学んだり、プログラミングの知識を吸収したりしながら、自分の為だけに貢献してくれる、マイ AI アシストを作り出していく。そうすれば当分 AI に職を奪われることはないだろう。

ビッグ・データがある場合には、統計的にみれば、正しい答えを AI は答えてくれるという間違った認識がある。しかし、それは間違いである。最近読んだ新井先生の著書に簡単な例でその事が記載されていた。
「私は、山口と岡山に行った。」
この文章を読んだ時に、人間であっても、山口が土地の名前なのか人の名前なのかはわからない。前後の文脈を理解する能力が必要になる。ネット上の全ての文章の、ビッグデータを解析したところで、山口が土地の名前か、人の名前かは確定できない。“この山口は人名”なのでクレンジングしようと思えるのは、人間だけで、その人の自己責任でクレンジングした結果、本来掘り下げるべきデータを除外して大発見を逃す事も十分あり得る。今回の例はあくまでも素人の筆者がクレンジングした結果なので、結果が異なっていても気にする必要は全く無い。様々な立場の研究者とモデルの違いを議論し、可能であれば実験結果、論文と、特許等からのデータ値を加味し、再構築して研究を加速して行くのが大事なのであろう。

Pirika マテリアル・ゲノム のページ
以下 PDF
第 1 回 イントロダクション 2018.8.23
第 2 回 データ収集と昔ながらのやり方 2018.8.24
第 0 回 物性推算と逆設計と呼んでいた時の話 2000.8.28
なんと 18 年前！
第 3a 回 ポリマー設計と 3 つの MI（その 1） 2018.9.3
第 3b 回 ポリマー設計と 3 つの MI（その 2） 2018.9.3
第 4a 回 MI に適した簡単なデータベースの利用法 2018.9.4
第 4b 回 複雑なポリマーのデータベース化 2018.9.7
プレゼン用：MI を使う時のデータベース構築法 2018.9.11
プレゼン用：複雑なポリマーの設計と DB 2018.9.15
第 5 回 データのクレンジング 2018.8.28
第 6a 回 ニューラルネットワーク法の初歩 2018.9.25
第 6b 回 ニューラルネットワーク法を使った Drug Design 2018.9.22
第 7 回 遺伝的アルゴリズム(GA)を理解しよう