SUPPORTING INFORMATION

A multi-task deep learning neural network for predicting flammability-related properties from molecular structures

Ao Yang[‡a,b], Yang Su[‡c,a], Zihao Wang[a], Saimeng Jin[a], Jingzheng Ren[b], Xiangping Zhang[d], Weifeng Shen[a] and James H. Clark[e]*

[a]*School of Chemistry and Chemical Engineering, Chongqing University, Chongqing 400044, China*

[b]*Department of Industrial and Systems Engineering, The Hong Kong Polytechnic University, Hong Kong SAR, China*

[c]*School of Intelligent Technology and Engineering, Chongqing University of Science and Technology, Chongqing 401331, China*

[d]*Beijing Key Laboratory of Ionic Liquids Clean Process, CAS Key Laboratory of Green Process and Engineering, Institute of Process Engineering, Chinese Academy of Sciences, Beijing, 100190, China*

[e]*Green Chemistry Centre of Excellence, University of York, York YO105D, UK*

**Correspondence should be addressed to: Weifeng Shen (shenweifeng@cqu.edu.cn)*

‡These authors contributed equally to this work.

S1. Signature descriptors

The signature descriptor[1] is an effective tool for encoding molecular structures with all atomic neighborhood into strings. This descriptor involving two forms: the first one is the atomic signatures generated at a specified root atom in a molecule, the other one is the molecular signatures combined from all atom signatures linearly. The size of each atomic signature is determined by a parameter named "height", h. An atomic Signature is a tree subgraph including all atoms/bonds extending out to the predefined distance h without backtracking. The atomic signature at a specified height h can be formally represented as ${}^{h}\sigma_{G(x)}$ for the root atom x of a 2D molecular graph $G = (V, E)$. Here, the symbols, V and E, refers to the vertex (atom) set and edge (bond) set, respectively. Acetaldoxime (CAS No. 107-29-9) is taken as an example to provide atomic signatures shown in **Figure S1**.

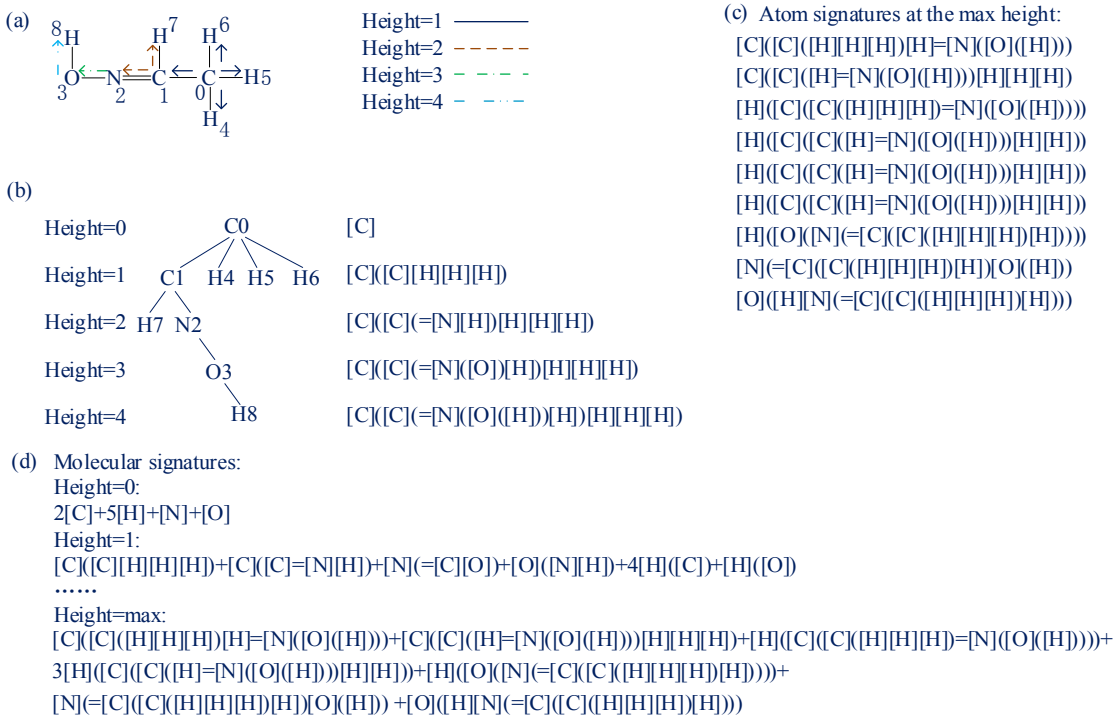

Figure S1. The translation example of signature descriptors for the molecule of acetaldoxime.

Note: The root atom numbered zero in the example for acetaldoxime is the terminal carbon in the recursive algorithm. An atomic signature begins from this root atom, steps forward a predetermined height, and records all the atoms encountered on the path are connected to the root atom. The process is repeated for all atoms in the molecule at a certain height, and then a molecular signature will be achieved by a linear combination of all the atomic signatures. In this figure, the molecular signature at the max height is exemplified for acetaldoxime: (a) the molecular structure; (b) tree form and atomic signatures of different heights from a certain atom; (c) the atomic signatures at the max height; (d) the molecular signature combined from the atomic signatures at the max height. Each atomic signature could be considered as a tree data structure stored in the format of DAG. In the encoding strings, the atoms found further down the tree of a branch point atom are marked by nested parenthesis. Single bonds between atoms are omitted in atomic signatures, while other bond-types are presented as follows ("="is a double bond; "#" is a triple bond; ":" is an aromatic bond.)

S2. Vectorizing vertexes of DAGs

As mentioned above, the signature descriptors are stored as DAGs in silico and represented by strings for easy reading, but it cannot be directly used like the other numerical descriptors (e.g. topological indices, occurrences of molecular sub-structures, microscopic and quantum properties) in the mathematical models. Hence, an approach was developed to transform atomic signatures into vectors for the subsequent computation. In this method, all bonds were traversed in a molecule and recorded into a list of strings. Afterwards, these strings like "A-B" (A and B represents atoms in a bond, "-" presents a bond type) were split into two part, "A" and "-B", as the samples to train a word embedding neural network with the skip-gram[2,3] algorithm. In this

way, the substrings "A" and "-B" can be mapped into vectors as inputs for each node in a Tree-LSTM[4] network. In other words, a molecule is considered as a sentence in the embedding algorithm, and "A" or "-B" is equivalent to a word.

The embedding neural network[2,5] is shown in **Figure S2**. In the training process, 170 substrings shown in **Table S1** are extracted from the chemical bonds of 23709 molecules (retrieved from PubChem[6]), and then they are converted to 170 vectors by the embedding algorithm. These 48-dimensional vectors will be used as input data for every node of the Tree-LSTM network.

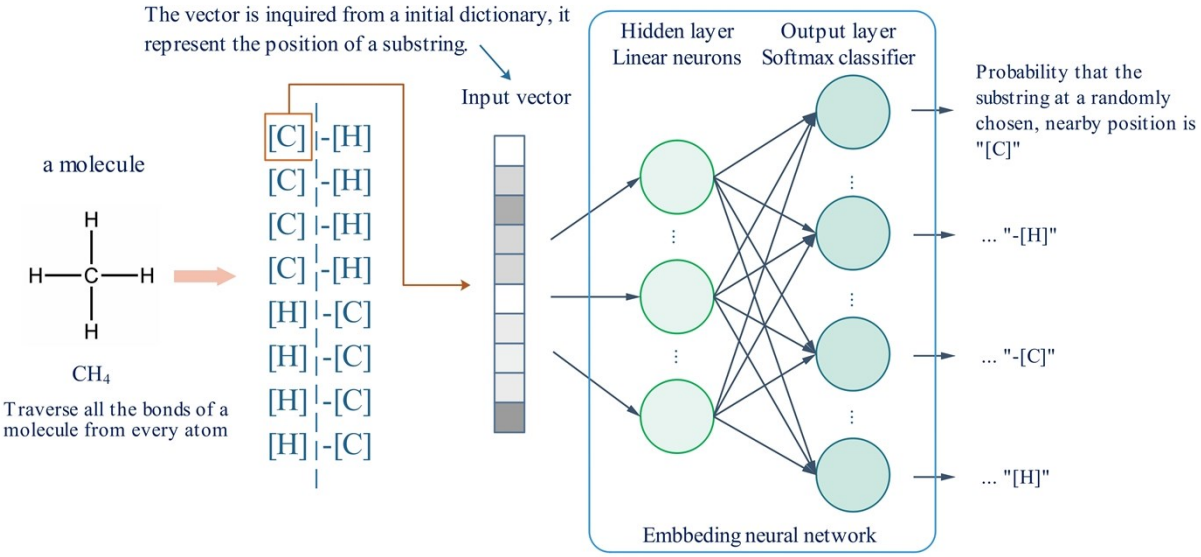

Figure S2. The embedding neural network for vectorizing the bond-strings.

Table S1. The list of substrings captured from chemical bonds

-[[C@@]|r,0]	=[C]	=[Al]	-[C|r,3]	-[n|r]	[P]	-[N|r,5]
[N]	-[N|r,0]	[C|r]	[O|r]	:[n|r,2]	-[O|r,1]	-[Cl+2]
[Se]	-[c|r]	[C]	[c|r]	[S]	-[[C@]|r,2]	-[Hg]
[Cl]	[P|r]	=[C|r]	[Pb]	[B]	=[Mg]	-[Fe]
-[O]	#[C-]	:[c|r,0]	-[[C@]|r]	-[P|r]	[O-]	[Ge]
-[3H]	-[[Si]|r]	=[S]	[Ga]	=[N|r]	=[N+]	[S-]
=[C|r,0]	-[N|r]	-[[C@]|r,3]	[Hg]	[F]	[I]	-[Tl]
-[S|r]	-[[C@@]|r,4]	[Sn]	-[B]	/=\[N]	-[[C@@]|r,2]	:[c|r,1]
-[N|r,3]	-[F]	-[[n+]|r]	-[Cl+]	-[[C@]|r,0]	:[n|r,0]	-[[Si]|r,0]
:[c|r,2]	:[c|r,5]	=[Ca]	:[c|r]	-[C|r,1]	=[N|r,1]	-[Al+]
:[s|r,1]	:[n|r,1]	-[S-]	=[B]	[s|r]	-[C|r,4]	-[C|r,2]
[Tl]	-[N|r,4]	-[S|r,1]	-[S]	[[Si]|r]	:[[n+]|r,0]	[O]
-[Cl]	/=\[C|r]	/=/[C|r]	/=\[C|r,0]	=[P|r]	-[V]	-[O|r,0]
-[Si]	=[Pb]	/=/[N]	#[C]	[N-]	#[N]	:[o|r,0]

/=/[C]	-[P]	#[B]	#[Si]	[Si+]	[2H]	[V]
-[Br]	:[o|r,1]	[Br]	-[Cl+3]	-[C@]	=[N]	=[O]
=[P]	-[c|r,2]	:[o|r]	-[[As]|r]	:[c|r,3]	-[O-]	=[C|r,1]
-[S|r,0]	-[P|r,0]	-[[C@]|r,4]	-[c|r,1]	[n|r]	-[c|r,3]	-[O|r,2]
-[N+]	-[C|r,0]	-[[C@@]|r,1]	[N|r]	-[I]	:[s|r]	-[Al]
[[n+]|r]	-[C@@]	:[c|r,6]	:[[n+]|r]	[Zn]	-[[n+]|r,2]	=[Fe]
-[Ge]	:[n|r]	-[[C@@]|r]	-[C|r]	[o|r]	-[H]	[N+]
-[C]	-[[C@@]|r,3]					

For example, the 48-dimensional vector representing the substring "=[C]" is shown as below:

[-1.578784935, -0.956886004, 0.513971263, 0.663698777, 0.683678891, 1.381063777, 0.116869597, 0.978596890, -0.044261189, 1.100085196, 0.460453439, 0.486235177, -0.229623520, 1.425614274, 1.613233129, 0.838176610, 0.757844079, -1.444321924, 0.391216566, 0.341012802, 0.281424870, 0.980665070, 0.519018905, -0.973128510, -1.507159523, 1.203545613, -0.988369549, 1.148227519, -0.461554489, 0.889267031, -0.196072090, -1.235736984, 0.380164882, 0.185392539, 0.035732845, -0.370580514, -0.107188924, -0.747444696, -0.693745025, 0.641271283, -0.050877589, -0.089852803, 0.638988558, -0.515607390, -1.068447242, 0.787464626, -0.122845660, 0.269591319]

S3. Tree-LSTM network

The Tree-LSTM neural network is employed for depicting molecular tree data-structures with the atomic signatures in this study. we chose the *N*-ary Tree-LSTM model[4] as listed from Eqs. S1 to S6 designed for the constituency tree to mimic the topological structure of an atomic signature. The gating vectors and memory cell updates of the Tree-LSTM are dependent on the states of child units, which is different than the standard LSTM. Especially, the *N*-ary Tree-LSTM model can be utilized in the tree structure where the branching factor is at most *N* and where children are ordered from 1 to *N*. For any node *j*, the hidden state and memory cell of its k^{th} child are written as h_{jk} and c_{jk}, respectively. The introduction of separate parameter matrices for each child *k* allows the N-ary Tree-LSTM model to learn more fine-grained conditioning on the states of a unit's children.

In this model, the variable x_j is the input vector representing a substring of a bond ("A" or "-B"), and the vector h_j is the output vector representing a molecular structure. The vector h_j is finally associated with the properties by the FNN. The FNN involves an input layer, a hidden

layer and an output layer. For other variables and functions in **Eqs.** (S1)-(S6), $W^{(i,o,u,f)}$, $U^{(i,o,u,f)}$, $b^{(i,o,u,f)}$ are parameters that need to be learned, and σ represents the activation function sigmoid. For example, the model can learn parameters $W^{(i)}$ such that the components of the input gate i_j have values close to 1 (*i.e.*, "open") when an important atom is given as input, and values close to 0 (*i.e.*, "closed") when the input is a less important atom.

$$i_j = \sigma(W^{(i)} x_j + \sum_{l=1}^{N} U_l^{(i)} h_{jl} + b^{(i)}) \quad \text{(S1)}$$

$$f_{jk} = \sigma(W^{(f)} x_j + \sum_{l=1}^{N} U_{kl}^{(f)} h_{jl} + b^{(f)}) \quad \text{(S2)}$$

$$o_j = \sigma(W^{(o)} x_j + \sum_{l=1}^{N} U_l^{(o)} h_{jl} + b^{(o)}) \quad \text{(S3)}$$

$$u_j = \tanh(W^{(u)} x_j + \sum_{l=1}^{N} U_l^{(u)} h_{jl} + b^{(u)} \quad \text{(S4)}$$

$$c_j = i_j \bullet u_j + \sum_{l=1}^{N} f_{jl} \bullet c_{jl} \quad \text{(S5)}$$

$$h_j = o_j \bullet \tanh(c_j) \quad \text{(S6)}$$

S4. Compound list

Table S2. The compounds correlated in the QSPR modeling for FPT

Name	Index
METHANE	0
ETHANE	1
PROPANE	2
ISOBUTANE	3
n-BUTANE	4
1,1,3-TRIMETHYLCYCLOHEXANE	5
n-PENTANE	6
ISOPENTANE	7
NEOPENTANE	8
n-HEXANE	9
2-METHYLPENTANE	10
2,2-DIMETHYLBUTANE	12
2,3-DIMETHYLBUTANE	13
n-HEPTANE	14
2-METHYLHEXANE	15
3-METHYLHEXANE	16
2,2-DIMETHYLPENTANE	18
2,4-DIMETHYLPENTANE	20
3,3-DIMETHYLPENTANE	21
2,2,3-TRIMETHYLBUTANE	22

n-OCTANE	23
2-METHYLHEPTANE	24
3-METHYLHEPTANE	25
4-METHYLHEPTANE	26
3-ETHYLHEXANE	27
2,2-DIMETHYLHEXANE	28
2,3-DIMETHYLHEXANE	29
2,4-DIMETHYLHEXANE	30
2,5-DIMETHYLHEXANE	31
3,4-DIMETHYLHEXANE	33
3-METHYL-3-ETHYLPENTANE	35
2,2,3-TRIMETHYLPENTANE	36
2,2,4-TRIMETHYLPENTANE	37
2,3,3-TRIMETHYLPENTANE	38
2,3,4-TRIMETHYLPENTANE	39
2,2,3,3-TETRAMETHYLBUTANE	40
n-NONANE	41
2,2,5-TRIMETHYLHEXANE	42
3,3,5-TRIMETHYLHEPTANE	43
2,4,4-TRIMETHYLHEXANE	44
3,3-DIETHYLPENTANE	45
2,2,3,3-TETRAMETHYLPENTANE	46
2,2,3,4-TETRAMETHYLPENTANE	47
2,2,4,4-TETRAMETHYLPENTANE	48
2,3,3,4-TETRAMETHYLPENTANE	49
SQUALANE	50
n-DECANE	51
iso-BUTYLCYCLOHEXANE	54
tert-BUTYLCYCLOHEXANE	55
n-UNDECANE	56
n-DODECANE	57
n-TRIDECANE	58
n-TETRADECANE	59
n-PENTADECANE	60
n-HEXADECANE	61
n-HEPTADECANE	62
n-OCTADECANE	63
n-NONADECANE	64
n-EICOSANE	66
3-METHYLNONANE	76
2-METHYLNONANE	77
4-METHYLNONANE	78
5-METHYLNONANE	79
2-METHYLOCTANE	81

3-METHYLOCTANE	82
4-METHYLOCTANE	83
3-ETHYLHEPTANE	84
2,2-DIMETHYLHEPTANE	85
CYCLOPENTANE	92
METHYLCYCLOPENTANE	93
ETHYLCYCLOPENTANE	94
n-PROPYLCYCLOPENTANE	100
CYCLOHEXANE	115
METHYLCYCLOHEXANE	116
ETHYLCYCLOHEXANE	117
1,1-DIMETHYLCYCLOHEXANE	118
cis-1,2-DIMETHYLCYCLOHEXANE	119
trans-1,2-DIMETHYLCYCLOHEXANE	120
cis-1,3-DIMETHYLCYCLOHEXANE	121
cis-1,4-DIMETHYLCYCLOHEXANE	123
1-trans-3,5-TRIMETHYLCYCLOHEXANE	125
ISOPROPYLCYCLOHEXANE	127
n-BUTYLCYCLOHEXANE	129
cis-DECAHYDRONAPHTHALENE	130
trans-DECAHYDRONAPHTHALENE	131
BICYCLOHEXYL	132
1,1-DIETHYLCYCLOHEXANE	133
CYCLOHEPTANE	135
CYCLOOCTANE	136
2,6-DIMETHYLHEPTANE	143
2,4-DIMETHYL-3-ETHYLPENTANE	145
ETHYLENE	146
PROPYLENE	147
1-BUTENE	149
cis-2-BUTENE	150
trans-2-BUTENE	151
ISOBUTENE	152
1-PENTENE	154
cis-2-PENTENE	155
trans-2-PENTENE	156
2-METHYL-1-BUTENE	157
3-METHYL-1-BUTENE	158
2-METHYL-2-BUTENE	159
1-HEXENE	161
cis-2-HEXENE	162
trans-2-HEXENE	163
trans-3-HEXENE	165
2-METHYL-1-PENTENE	166

3-METHYL-1-PENTENE	167
4-METHYL-1-PENTENE	168
2-METHYL-2-PENTENE	169
3-METHYL-cis-2-PENTENE	170
4-METHYL-1-HEXENE	171
4-METHYL-cis-2-PENTENE	172
4-METHYL-trans-2-PENTENE	173
2-ETHYL-1-BUTENE	174
2,3-DIMETHYL-1-BUTENE	175
3,3-DIMETHYL-1-BUTENE	176
2,3-DIMETHYL-2-BUTENE	177
1-HEPTENE	179
cis-2-HEPTENE	180
trans-2-HEPTENE	181
trans-3-HEPTENE	182
2-METHYL-1-HEXENE	183
3-METHYL-1-HEXENE	185
trans-2-NONENE	190
VINYLCYCLOHEXANE	191
2,3,3-TRIMETHYL-1-BUTENE	193
1-OCTENE	195
trans-2-OCTENE	196
2,4,4-TRIMETHYL-1-PENTENE	197
2,4,4-TRIMETHYL-2-PENTENE	198
1-NONENE	200
1-DECENE	201
1-UNDECENE	202
1-DODECENE	203
1-TRIDECENE	204
1-TETRADECENE	205
1-HEXADECENE	207
1-OCTADECENE	208
CYCLOPENTENE	210
CYCLOHEXENE	211
CYCLOHEPTENE	214
CYCLOOCTENE	215
cis-2-OCTENE	216
trans-3-OCTENE	217
cis-4-OCTENE	218
trans-4-OCTENE	219
cis-3-OCTENE	220
VINYLCYCLOHEXENE	224
1-METHYLCYCLOPENTENE	225
2,3-DIMETHYL-1-HEXENE	228

d-LIMONENE	229
TERPINOLENE	230
1,3-BUTADIENE	238
cis-1,3-PENTADIENE	240
trans-1,3-PENTADIENE	241
1,4-PENTADIENE	242
ISOPRENE	244
1,5-HEXADIENE	245
3-METHYL-1,2-BUTADIENE	246
METHYLCYCLOPENTADIENE	247
1,4-HEXADIENE	248
trans,trans-2,4-HEXADIENE	249
DICYCLOPENTADIENE	251
alpha-PHELLANDRENE	252
beta-PHELLANDRENE	253
2,3-DIMETHYL-1,3-BUTADIENE	254
cis,trans-2,4-HEXADIENE	255
1,5,9-CYCLODODECATRIENE	257
2,5-DIMETHYL-1,5-HEXADIENE	258
2,5-DIMETHYL-2,4-HEXADIENE	259
1,3-CYCLOHEXADIENE	260
1,4-CYCLOHEXADIENE	261
1,5-CYCLOOCTADIENE	262
trans-1,3-HEXADIENE	263
trans-2-METHYL-1,3-PENTADIENE	264
1,9-DECADIENE	265
1-PENTYNE	271
2-HEXYNE	273
2-PENTYNE	275
1-HEXYNE	276
1-OCTYNE	278
1-PENTENE-3-YNE	281
1-NONYNE	284
1-DECYNE	285
BENZENE	286
TOLUENE	287
ETHYLBENZENE	289
o-XYLENE	290
m-XYLENE	291
p-XYLENE	292
n-PROPYLBENZENE	294
CUMENE	295
o-ETHYLTOLUENE	296
m-ETHYLTOLUENE	297

p-ETHYLTOLUENE	298
1,2,3-TRIMETHYLBENZENE	299
1,2,4-TRIMETHYLBENZENE	300
MESITYLENE	301
n-BUTYLBENZENE	303
ISOBUTYLBENZENE	304
sec-BUTYLBENZENE	305
tert-BUTYLBENZENE	306
o-CYMENE	307
m-CYMENE	308
p-CYMENE	309
o-DIETHYLBENZENE	310
m-DIETHYLBENZENE	311
p-DIETHYLBENZENE	312
1,2,3,4-TETRAMETHYLBENZENE	315
1,2,3,5-TETRAMETHYLBENZENE	316
1,2,4,5-TETRAMETHYLBENZENE	317
1,3,5-TRIISOPROPYLBENZENE	322
PENTAMETHYLBENZENE	325
m-DIISOPROPYLBENZENE	327
p-DIISOPROPYLBENZENE	328
1,2,4-TRIETHYLBENZENE	329
n-HEPTYLBENZENE	332
n-DECYLBENZENE	334
CYCLOHEXYLBENZENE	337
BIPHENYL	338
p-TERPHENYL	339
m-TERPHENYL	340
o-TERPHENYL	341
1,1-DIPHENYLETHANE	342
DIPHENYLMETHANE	343
1,2-DIPHENYLETHANE	344
n-PENTYLBENZENE	347
n-HEXYLBENZENE	348
n-OCTYLBENZENE	349
n-NONYLBENZENE	350
n-DODECYLBENZENE	354
2-PHENYLBUTENE-1	362
STYRENE	378
o-METHYLSTYRENE	379
m-METHYLSTYRENE	380
5-o-TOLYL-2-PENTENE	381
m-ETHYLSTYRENE	383
p-METHYLSTYRENE	385

alpha-METHYLSTYRENE	386
m-DIVINYLBENZENE	387
ETHYNYLBENZENE	388
cis-1-PROPENYLBENZENE	391
trans-1-PROPENYLBENZENE	392
NAPHTHALENE	395
1-METHYLNAPHTHALENE	396
2-METHYLNAPHTHALENE	397
1-ETHYLNAPHTHALENE	398
1,2,3,4-TETRAHYDRONAPHTHALENE	399
1-n-DECYLNAPHTHALENE	403
2-ETHYLNAPHTHALENE	410
1,2,3-TRIMETHYLINDENE	413
METHYLCYCLOPENTADIENE DIMER	414
INDENE	423
ANTHRACENE	424
PHENANTHRENE	425
PYRENE	427
VINYLNORBORNENE	431
1,3-DIMETHYLADAMANTANE	434
INDANE	438
alpha-TERPINENE	439
gamma-TERPINENE	440
2-NORBORNENE	441
5-ETHYLIDENE-2-NORBORNENE	442
sec-BUTYLCYCLOHEXANE	445
CAMPHENE	455
alpha-PINENE	456
beta-PINENE	457
ACETALDEHYDE	475
PROPANAL	476
1,2,3,6-TETRAHYDROBENZALDEHYDE	477
BUTANAL	478
2-METHYLPROPANAL	479
PENTANAL	480
HEPTANAL	481
HEXANAL	482
OCTANAL	483
NONANAL	484
CYCLOHEXANECARBOXALDEHYDE	485
2-ETHYLHEXANAL	486
2-METHYL-2-PENTENAL	490
2-ETHYL-2-HEXENAL	491
DECANAL	492

UNDECANAL	493
DODECANAL	494
2-METHYLBUTYRALDEHYDE	496
3-METHYLBUTYRALDEHYDE	497
ACROLEIN	499
trans-CROTONALDEHYDE	501
METHACROLEIN	502
2-PHENYLPROPIONALDEHYDE	503
o-TOLUALDEHYDE	504
p-TOLUALDEHYDE	505
BENZALDEHYDE	506
SALICYLALDEHYDE	507
m-TOLUALDEHYDE	510
PARALDEHYDE	513
ACETONE	514
METHYL ETHYL KETONE	515
3-PENTANONE	516
METHYL ISOBUTYL KETONE	517
3-METHYL-2-PENTANONE	518
5-HEXEN-2-ONE	519
3-HEPTANONE	520
4-HEPTANONE	521
3-HEXANONE	522
2-PENTANONE	523
METHYL ISOPROPYL KETONE	524
2-HEXANONE	525
2-HEPTANONE	526
5-METHYL-2-HEXANONE	527
MESITYL OXIDE	528
3,3-DIMETHYL-2-BUTANONE	529
DIISOBUTYL KETONE	530
DIISOPROPYL KETONE	531
2-PYRROLIDONE	532
N-METHYL-2-PYRROLIDONE	533
5-NONANONE	535
2-NONANONE	536
ANTHRAQUINONE	537
ACETYLACETONE	538
ISOPHORONE	539
CYCLOPENTANONE	540
CYCLOHEXANONE	541
2-OCTANONE	542
ACETOPHENONE	544
beta-PROPIOLACTONE	545

gamma-BUTYROLACTONE	546
epsilon-CAPROLACTONE	547
gamma-VALEROLACTONE	548
ETHYL ISOPROPYL KETONE	549
METHYL ISOPROPENYL KETONE	550
2-CYCLOHEXYL CYCLOHEXANONE	551
QUINONE	552
DIKETENE	553
METHANOL	555
ETHANOL	556
1-PROPANOL	557
ISOPROPANOL	558
1-BUTANOL	559
2-METHYL-1-PROPANOL	560
2-BUTANOL	561
2-METHYL-2-PROPANOL	562
1-PENTANOL	563
2-PENTANOL	564
2-METHYL-2-BUTANOL	565
2-METHYL-1-BUTANOL	566
2,2-DIMETHYL-1-PROPANOL	567
1-HEXANOL	568
2-HEXANOL	569
3-HEXANOL	570
2-METHYL-1-PENTANOL	571
3-METHYL-1-PENTANOL	572
3-METHYL-3-PENTANOL	573
3-PENTANOL	574
2-ETHYL-1-HEXANOL	575
2-METHYL-1-HEXANOL	576
3-METHYL-1-BUTANOL	577
3-METHYL-2-BUTANOL	578
1-HEPTANOL	579
2-HEPTANOL	580
3-HEPTANOL	581
5-METHYL-1-HEXANOL	582
4-METHYL-2-PENTANOL	583
1-OCTANOL	585
2-OCTANOL	586
1-NONANOL	587
2-NONANOL	588
1-DECANOL	589
1-UNDECANOL	590
8-METHYL-1-NONANOL	591

1-DODECANOL	592
1-TRIDECANOL	593
1-TETRADECANOL	594
1-HEXADECANOL	596
1-HEPTADECANOL	597
1-OCTADECANOL	598
2-ETHYL-1-BUTANOL	599
CYCLOHEXANOL	603
1-METHYLCYCLOHEXANOL	604
cis-2-METHYLCYCLOHEXANOL	605
trans-2-METHYLCYCLOHEXANOL	606
cis-3-METHYLCYCLOHEXANOL	607
trans-3-METHYLCYCLOHEXANOL	608
cis-4-METHYLCYCLOHEXANOL	609
L-MENTHOL	611
alpha-TERPINEOL	614
TETRAHYDROFURFURYL ALCOHOL	618
ALLYL ALCOHOL	619
2-PHENYL-2-PROPANOL	620
2,3-XYLENOL	622
2,4-XYLENOL	624
2,5-XYLENOL	626
2,6-XYLENOL	628
PROPARGYL ALCOHOL	631
BENZYL ALCOHOL	632
PHENOL	633
o-CRESOL	634
m-CRESOL	635
p-CRESOL	636
o-ETHYLPHENOL	637
p-HYDROQUINONE	638
p-ETHYLPHENOL	639
p-tert-BUTYLPHENOL	640
BISPHENOL A	641
NONYLPHENOL	642
ETHYLENE GLYCOL	643
DIETHYLENE GLYCOL	644
TRIETHYLENE GLYCOL	645
TETRAETHYLENE GLYCOL	646
1,2-PROPYLENE GLYCOL	653
1,3-PROPYLENE GLYCOL	654
DIPROPYLENE GLYCOL	655
NEOPENTYL GLYCOL	656
2-BUTYNE-1,4-DIOL	657

TRIPROPYLENE GLYCOL	658
2-METHYL-1,3-PROPANEDIOL	660
1,2-BUTANEDIOL	661
1,3-BUTANEDIOL	662
HEXYLENE GLYCOL	663
GLYCEROL	665
p-tert-BUTYLCATECHOL	669
2,2,4-TRIMETHYL-1,3-PENTANEDIOL	670
2-METHYL-1,3-PENTANEDIOL	671
2,3-BUTANEDIOL	672
cis-2-BUTENE-1,4-DIOL	673
trans-2-BUTENE-1,4-DIOL	674
1,4-BUTANEDIOL	675
1,5-PENTANEDIOL	676
1,6-HEXANEDIOL	677
1,2-BENZENEDIOL	678
1,3-BENZENEDIOL	679
PENTAERYTHRITOL	680
TRIMETHYLOLPROPANE	681
FORMIC ACID	685
ACETIC ACID	686
PROPIONIC ACID	687
n-DECANOIC ACID	688
n-BUTYRIC ACID	690
n-PENTANOIC ACID	692
n-NONANOIC ACID	693
ISOBUTYRIC ACID	694
ISOVALERIC ACID	695
n-HEXANOIC ACID	696
2-METHYLHEXANOIC ACID	697
n-OCTANOIC ACID	699
CYCLOPENTYLACETIC ACID	701
n-DODECANOIC ACID	703
trans-CROTONIC ACID	708
ACRYLIC ACID	711
METHACRYLIC ACID	712
OLEIC ACID	713
BENZOIC ACID	715
SALICYLIC ACID	718
ADIPIC ACID	719
PHTHALIC ACID	721
TEREPHTHALIC ACID	723
PERACETIC ACID	724
ACETIC ANHYDRIDE	725

PROPIONIC ANHYDRIDE	726
BUTYRIC ANHYDRIDE	727
GLUTARIC ANHYDRIDE	730
PHTHALIC ANHYDRIDE	731
MALEIC ANHYDRIDE	732
TRIMELLITIC ANHYDRIDE	733
METHYL FORMATE	734
ETHYL FORMATE	735
n-PROPYL FORMATE	736
n-BUTYL FORMATE	737
ISOBUTYL FORMATE	738
n-PENTYL FORMATE	739
n-OCTYL FORMATE	740
METHYL ACETATE	744
ETHYL ACETATE	745
n-PROPYL ACETATE	746
n-BUTYL ACETATE	747
ISOBUTYL ACETATE	748
ISOPENTYL ACETATE	749
ALLYL ACETATE	750
ISOPROPYL ACETATE	751
sec-BUTYL ACETATE	752
VINYL ACETATE	753
METHYL PROPIONATE	754
ETHYL PROPIONATE	755
n-PROPYL PROPIONATE	756
n-BUTYL PROPIONATE	757
n-PROPYL n-BUTYRATE	758
VINYL PROPIONATE	759
METHYL n-BUTYRATE	760
ETHYL n-BUTYRATE	761
ETHYL CAPRYLATE	763
METHYL ACRYLATE	764
ETHYL ACRYLATE	765
n-BUTYL ACRYLATE	767
ETHYL ISOVALERATE	770
METHYL METHACRYLATE	771
ETHYL METHACRYLATE	772
DIOCTYL PHTHALATE	774
DIISOOCTYL PHTHALATE	775
n-PENTYL ACETATE	777
2-ETHYLHEXYL ACETATE	778
BENZYL ACETATE	779
ISOBUTYL ISOBUTYRATE	780

ISOPENTYL ISOVALERATE	781
n-HEXYL ACETATE	783
BENZYL BENZOATE	784
n-BUTYL BENZOATE	785
ETHYLENE CARBONATE	786
n-HEPTYL ACETATE	787
n-OCTYL ACETATE	788
n-NONYL ACETATE	789
n-DECYL ACETATE	790
DIISODECYL PHTHALATE	791
METHYL SALICYLATE	793
DI-n-NONYL PHTHALATE	794
DIISOBUTYL PHTHALATE	796
DIMETHYL ISOPHTHALATE	797
DI-n-DECYL PHTHALATE	798
DI-n-HEXYL PHTHALATE	800
DIMETHYL TEREPHTHALATE	801
n-BUTYL STEARATE	803
DIBUTYL SEBACATE	804
n-BUTYL n-BUTYRATE	805
2-ETHYLHEXYL ACRYLATE	806
ETHYLENE GLYCOL DIACETATE	807
ISOBUTYL METHACRYLATE	808
n-BUTYL METHACRYLATE	809
METHYL BENZOATE	810
ETHYL BENZOATE	811
DIETHYL CARBONATE	812
DIETHYL OXALATE	813
DIETHYL MALONATE	814
METHYL DECANOATE	815
ISOPROPYL MYRISTATE	816
DIMETHYL-1,4-CYCLOHEXANEDICARBOXYLATE	817
2-METHYLTETRAHYDROFURAN	820
DIMETHYL ETHER	821
DIETHYL ETHER	822
DIISOPROPYL ETHER	823
DI-n-BUTYL ETHER	824
METHYL tert-BUTYL ETHER	825
DI-sec-BUTYL ETHER	826
METHYL ETHYL ETHER	827
DI-n-HEXYL ETHER	832
METHYL n-BUTYL ETHER	833
DIVINYL ETHER	834
ETHYL PROPYL ETHER	835

ANETHOLE	838
1,4-DIOXANE	839
TRIOXANE	840
DI-tert-BUTYL ETHER	841
DI-n-OCTYL ETHER	842
DI-n-PENTYL ETHER	843
METHYL sec-BUTYL ETHER	844
METHYL tert-PENTYL ETHER	845
tert-BUTYL ETHYL ETHER	846
METHYLAL	849
ACETAL	850
TRIETHYL ORTHOFORMATE	854
ETHYL HYDROPEROXIDE	856
1,2-EPOXY-2-METHYLPROPANE	858
ETHYLENE OXIDE	859
1,2-PROPYLENE OXIDE	860
ETHYL VINYL ETHER	862
DI-n-PROPYL ETHER	863
BUTYL VINYL ETHER	864
n-BUTYL ETHYL ETHER	865
BIS(alpha-METHYLBENZYL) ETHER	866
ALLYL ETHER	870
TRIETHYLENE GLYCOL DIMETHYL ETHER	871
1,2-DIMETHOXYETHANE	872
DIETHYLENE GLYCOL DIMETHYL ETHER	873
TETRAETHYLENE GLYCOL DIMETHYL ETHER	874
DIETHYLENE GLYCOL DIETHYL ETHER	875
DIETHYLENE GLYCOL DI-n-BUTYL ETHER	876
ANISOLE	878
PHENETOLE	879
DIBENZYL ETHER	880
ALLYL GLYCIDYL ETHER	881
DIPHENYL ETHER	882
1,1-DIMETHOXYETHANE	885
METHYL VINYL ETHER	887
1,2-EPOXYBUTANE	888
CUMENE HYDROPEROXIDE	889
t-BUTYL HYDROPEROXIDE	890
DICUMYL PEROXIDE	892
2,5-DIHYDROFURAN	894
FURAN	895
TETRAHYDROFURAN	896
DIBENZOFURAN	897
DI-t-BUTYL PEROXIDE	899

2-METHYLBENZOFURAN	902
1,2-DIMETHOXYPROPANE	906
2-METHYL-1,3-DIOXOLANE	909
2,5-DIMETHYLTETRAHYDROFURAN	910
METHYL CHLORIDE	912
ETHYL CHLORIDE	913
VINYL CHLORIDE	914
1,4-DICHLORO-trans-2-BUTENE	915
1,4-DICHLOROBUTANE	918
1,5-DICHLOROPENTANE	919
1,1-DICHLOROETHANE	924
1,2-DICHLOROETHANE	925
1,2-DICHLOROPROPANE	928
ISOPROPYL CHLORIDE	932
PROPARGYL CHLORIDE	933
1,2,3-TRICHLOROPROPANE	934
ISOBUTYL CHLORIDE	935
tert-BUTYL CHLORIDE	936
1,2-DICHLOROBUTANE	937
2,3-DICHLOROBUTANE	938
TRICHLOROETHYLENE	939
3-CHLOROPROPENE	941
CHLOROCYCLOHEXANE	942
BENZYL CHLORIDE	944
3-CHLORO-o-XYLENE	945
4-CHLORO-o-XYLENE	946
MONOCHLOROBENZENE	947
o-DICHLOROBENZENE	948
m-DICHLOROBENZENE	949
p-DICHLOROBENZENE	950
HEXACHLOROBENZENE	951
o-CHLOROTOLUENE	953
p-CHLOROTOLUENE	954
2,4-DICHLOROTOLUENE	955
cis-1,2-DICHLOROETHYLENE	956
trans-1,2-DICHLOROETHYLENE	957
CHLOROPRENE	959
n-BUTYL CHLORIDE	961
sec-BUTYL CHLORIDE	962
1-CHLOROPENTANE	963
1-CHLORONAPHTHALENE	964
1,1-DICHLOROETHYLENE	966
1,2,4-TRICHLOROBENZENE	967
1,4-DICHLORO-cis-2-BUTENE	968

2-CHLOROPROPENE	970
m-CHLOROBENZOYL CHLORIDE	971
3,4-DICHLORO-1-BUTENE	972
(Z)-1,3-DICHLORO-2-BUTENE	973
BENZYL DICHLORIDE	974
1,3-DIFLUOROBENZENE	1004
1,2-DIFLUOROBENZENE	1005
BROMOETHANE	1008
1-BROMOBUTANE	1015
1-BROMOHEPTANE	1016
m-DIBROMOBENZENE	1019
BROMOBENZENE	1021
n-PROPYL IODIDE	1024
ISOPROPYL IODIDE	1025
1,4-DIFLUOROBENZENE	1026
IODOBENZENE	1027
METHYLAMINE	1036
DIMETHYLAMINE	1037
TRIMETHYLAMINE	1038
2-METHYL-2-AMINOBUTANE	1040
TRIETHYLAMINE	1041
DI-n-PROPYLAMINE	1042
N,N-DIMETHYL-n-BUTYLAMINE	1043
N,N-DIETHYLMETHYLAMINE	1044
DIETHYLAMINE	1045
n-PROPYLAMINE	1046
n-BUTYLAMINE	1047
n-PENTYLAMINE	1048
ISOBUTYLAMINE	1049
CYCLOPROPYLAMINE	1050
HYDRAZINE	1051
DIISOBUTYLAMINE	1052
ISOPROPYLAMINE	1053
PYRROLE	1055
METHYL DIETHANOLAMINE	1056
MONOETHANOLAMINE	1057
DIETHANOLAMINE	1058
TRIETHANOLAMINE	1059
sec-BUTYLAMINE	1060
tert-BUTYLAMINE	1061
CYCLOHEXYLAMINE	1062
DEHYDROABIETYLAMINE	1063
HEXAMETHYLENEDIAMINE	1064
BENZYLAMINE	1066

TRIETHYLENEDIAMINE	1067
o-TOLUIDINE	1069
m-TOLUIDINE	1070
p-TOLUIDINE	1071
TRIETHYLENETETRAMINE	1072
ALLYLAMINE	1073
ETHYLENEDIAMINE	1074
ETHYLENEIMINE	1075
DIISOPROPYLAMINE	1076
DI-n-BUTYLAMINE	1077
PIPERIDINE	1078
p-AMINODIPHENYLAMINE	1080
QUINOLINE	1081
N-AMINOETHYL PIPERAZINE	1083
HEXAMETHYLENETETRAMINE	1084
1,2-PROPANEDIAMINE	1085
N,N-DIETHYLANILINE	1086
N-METHYLPYRROLE	1087
PYRAZINE	1088
DIPHENYLAMINE	1089
PHENYLHYDRAZINE	1090
QUINALDINE	1092
NITROMETHANE	1093
NITROETHANE	1094
1-NITROPROPANE	1095
2-NITROPROPANE	1096
HYDRACRYLONITRILE	1097
MORPHOLINE	1098
PYRROLIDINE	1099
N-METHYLPYRROLIDINE	1100
1-NITROBUTANE	1102
HYDROGEN CYANIDE	1103
ACETONITRILE	1104
PROPIONITRILE	1105
ACRYLONITRILE	1106
METHACRYLONITRILE	1107
SUCCINONITRILE	1108
ADIPONITRILE	1109
o-NITROTOLUENE	1110
p-NITROTOLUENE	1111
m-NITROTOLUENE	1112
BUTYRONITRILE	1114
VALERONITRILE	1115
MALONONITRILE	1116

HEXANENITRILE	1117
ISOBUTYRONITRILE	1118
BENZONITRILE	1120
PYRIDINE	1121
ANILINE	1122
2,4-TOLUENE DIISOCYANATE	1123
HEXAMETHYLENEIMINE	1124
N-METHYLANILINE	1125
N,N-DIMETHYLANILINE	1126
2-METHYLPYRIDINE	1127
cis-CROTONITRILE	1128
METHYL MERCAPTAN	1130
n-PROPYL MERCAPTAN	1132
tert-BUTYL MERCAPTAN	1133
ISOBUTYL MERCAPTAN	1134
sec-BUTYL MERCAPTAN	1135
n-HEXYL MERCAPTAN	1136
n-NONYL MERCAPTAN	1137
n-OCTYL MERCAPTAN	1138
ISOPROPYL MERCAPTAN	1139
CYCLOHEXYL MERCAPTAN	1140
BENZYL MERCAPTAN	1141
METHYL ETHYL SULFIDE	1142
METHYL t-BUTYL SULFIDE	1144
DI-n-PROPYL SULFIDE	1146
DIETHYL SULFIDE	1147
DIMETHYL SULFIDE	1149
THIOPHENE	1150
BENZOTHIOPHENE	1151
DIETHYL DISULFIDE	1153
n-DECYL MERCAPTAN	1155
n-PENTYL MERCAPTAN	1156
DIMETHYL DISULFIDE	1157
DI-n-PROPYL DISULFIDE	1158
n-DODECYL MERCAPTAN	1166
tert-OCTYL MERCAPTAN	1167
n-HEPTYL MERCAPTAN	1168
DIALLYL SULFIDE	1169
n-BUTYL MERCAPTAN	1170
PHENYL MERCAPTAN	1171
TETRAHYDROTHIOPHENE	1172
DIMETHYL SULFOXIDE	1173
SULFOLANE	1174
DI-n-PROPYL SULFONE	1176

DI-n-BUTYL SULFONE	1177
ACETYL CHLORIDE	1179
CHLOROACETIC ACID	1180
DICHLOROACETYL CHLORIDE	1182
BENZOYL CHLORIDE	1184
p-CHLOROBENZOTRIFLUORIDE	1185
2,4-DICHLOROBENZOTRIFLUORIDE	1186
o-CHLOROANILINE	1187
FLUOROBENZENE	1188
CHLOROMETHYL METHYL ETHER	1189
3-CHLORO-1,2-PROPANEDIOL	1190
1,3-DICHLORO-2-PROPANOL	1191
HEXAFLUOROBENZENE	1192
DI(2-CHLOROETHYL)ETHER	1194
3-(METHYLMERCAPTO)PROPANAL	1196
3-MERCAPTOPROPIONIC ACID	1201
N,N-DIMETHYLFORMAMIDE	1204
epsilon-CAPROLACTAM	1208
alpha-EPICHLOROHYDRIN	1209
ACETONE CYANOHYDRIN	1210
TRI-n-BUTYL BORATE	1211
TRIPHENYLPHOSPHINE	1212
HEXAMETHYL PHOSPHORAMIDE	1213
NITROBENZENE	1214
ACETALDOL	1216
FURFURAL	1217
PROPYLENE GLYCOL MONOMETHYL ETHER ACETATE	1224
DIPROPYLENE GLYCOL MONOMETHYL ETHER ACETATE	1225
METHOXYDIHYDROPYRAN	1226
CARBON DISULFIDE	1252
HEXAMETHYLDISILAZANE	1262
HEXAMETHYLDISILOXANE	1263
HEXAMETHYLCYCLOTRISILOXANE	1264
DODECAMETHYLPENTASILOXANE	1267
METHYL VINYL DICHLOROSILANE	1269
3-(TRIMETHOXYSILYL)-1-PROPANETHIOL	1273
TETRAMETHYLSILANE	1278
OCTAMETHYLCYCLOTETRASILOXANE	1281
3-OCTANONE	1295
3-NONANONE	1297
2-CYCLOHEXENE-1-ONE	1300
2,6-DIMETHYLOCTANE	1305
2,7-DIMETHYLOCTANE	1306
m-ETHYLPHENOL	1307

2,6-DI-tert-BUTYL-p-CRESOL	1308
2-PHENYLETHANOL	1309
2,6-DIMETHYL-4-HEPTANOL	1310
1-PHENYL-1-PROPANOL	1311
1-PHENYL-2-PROPANOL	1312
2-PHENYL-1-PROPANOL	1313
3-PHENYL-1-PROPANOL	1314
alpha-METHYLBENZYL ALCOHOL	1316
o-TOLUALCOHOL	1318
m-TOLUALCOHOL	1319
p-tert-AMYLPHENOL	1321
p-CUMYLPHENOL	1322
2,4-PENTANEDIOL	1324
3-METHYL-trans-2-PENTENE	1325
5-METHYL-1-HEXENE	1326
1,4-CYCLOHEXANEDIMETHANOL	1328
2-METHYL-1-OCTENE	1330
2-METHYL-1-HEPTENE	1331
NEOPENTANOIC ACID	1337
2-ETHYL HEXANOIC ACID	1339
n-HEPTANOIC ACID	1340
NEOHEXANOIC ACID	1341
2-ETHYL BUTYRIC ACID	1355
METHYL MALEIC ANHYDRIDE	1362
tert-BUTYL FORMATE	1364
n-HEXYL FORMATE	1365
1,2-HEXADIENE	1367
CYCLOHEXYL FORMATE	1368
tert-BUTYL ACETATE	1369
METHYL ISOBUTYRATE	1370
ETHYL ISOBUTYRATE	1371
BENZYL FORMATE	1373
ALLYL METHACRYLATE	1375
CYCLOHEXYL ACETATE	1377
PROPYLENE CARBONATE	1378
GLYCERYL TRIACETATE	1379
DIETHYL PHTHALATE	1382
DI-n-BUTYL PHTHALATE	1383
DIMETHYL PHTHALATE	1384
DIETHYL SUCCINATE	1385
DIHEXYL ADIPATE	1386
ETHYLIDENE DIACETATE	1387
DIALLYL MALEATE	1388
DIBUTYL MALEATE	1389

ISOBUTYL ACRYLATE	1390
DIETHYL MALEATE	1392
DIMETHYL MALEATE	1393
VINYL PIVALATE	1396
DIMETHYL CARBONATE	1397
n-PROPYL BENZOATE	1398
DIISOBUTYL ETHER	1399
1-HEPTYNE	1400
1,3-DIOXANE	1403
ETHYLAL	1404
1,2-DIETHOXYETHANE	1405
2,3-DIHYDROFURAN	1406
1,1-DICHLOROPROPANE	1408
1,3-DICHLOROPROPANE	1409
2,3-DICHLOROPROPENE	1410
1,2,3-TRICHLOROBENZENE	1415
1,3,5-TRICHLOROBENZENE	1416
BENZOTRIFLUORIDE	1430
2-BROMOBUTANE	1432
p-BROMOTOLUENE	1452
n-HEXYL IODIDE	1453
n-BUTYL IODIDE	1454
CHLOROTRIFLUOROETHYLENE	1459
n-HEXYLAMINE	1465
n-HEPTYLAMINE	1466
n-OCTYLAMINE	1467
n-NONYLAMINE	1468
n-DECYLAMINE	1469
n-DODECYLAMINE	1470
N-ETHYL-2-METHYLALLYLAMINE	1471
PENTAETHYLENE HEXAMINE	1472
TRIALLYLAMINE	1473
TRI-n-BUTYLAMINE	1474
DIETHYLENETRIAMINE	1475
TETRAETHYLENEPENTAMINE	1476
TRIPROPYLAMINE	1477
VINYLACETONITRILE	1478
n-BUTYL ISOCYANATE	1480
CYCLOHEXYL ISOCYANATE	1481
o-ETHYLANILINE	1482
o-PHENYLENEDIAMINE	1483
PROPYLENEIMINE	1484
m-PHENYLENEDIAMINE	1485
CYCLOPENTYLAMINE	1487

DICYCLOHEXYLAMINE	1488
N-METHYLCYCLOHEXYLAMINE	1489
N-AMINOETHYL ETHANOLAMINE	1490
DIPHENYLMETHANE-4,4'-DIISOCYANATE	1494
N,N'-DIPHENYL-p-PHENYLENEDIAMINE	1495
DIALLYLAMINE	1497
m-DINITROBENZENE	1498
o-DINITROBENZENE	1499
2,4-DINITROTOLUENE	1501
3,4-DINITROTOLUENE	1503
p-PHENYLENEDIAMINE	1508
PHENYL ISOCYANATE	1509
PIPERAZINE	1510
PYRIDAZINE	1511
PYRIMIDINE	1512
PYRAZOLE	1513
o-NITROANILINE	1520
p-NITROANILINE	1521
m-NITROANILINE	1522
ISOQUINOLINE	1525
p-AMINODIPHENYL	1527
NICOTINONITRILE	1530
2,6-DIETHYLANILINE	1531
METHYL ISOCYANATE	1532
2,4,6-TRIMETHYLPYRIDINE	1534
2,6-DIMETHYLPYRIDINE	1535
3-METHYLPYRIDINE	1536
METHYLGLUTARONITRILE	1537
4-METHYLPYRIDINE	1538
1,6-HEXAMETHYLENE DIISOCYANATE	1539
ISOPHORONE DIISOCYANATE	1540
THIACYCLOPROPANE	1541
METHYL n-BUTYL SULFIDE	1542
ETHYL t-BUTYL SULFIDE	1543
TRIMETHYLENE SULFIDE	1544
tert-DODECYL MERCAPTAN	1546
tert-NONYL MERCAPTAN	1547
2-METHYLTHIOPHENE	1548
3-METHYLTHIOPHENE	1549
CAMPHOR	1550
FORMAMIDE	1551
N-METHYLFORMAMIDE	1552
DIACETONE ALCOHOL	1554
FURFURYL ALCOHOL	1555

N,N-DIMETHYLACETAMIDE	1556
p-METHOXYPHENOL	1559
2-METHOXYETHANOL	1560
2-ETHOXYETHANOL	1561
2-BUTOXYETHANOL	1562
2-(2-METHOXYETHOXY)ETHANOL	1563
2-(2-ETHOXYETHOXY)ETHANOL	1564
2-AMINOETHOXYETHANOL	1565
ETHYLCHLOROACETATE	1566
ACETALDOXIME	1567
TRIETHYLENE GLYCOL MONOMETHYL ETHER	1569
2-HYDROXYETHYL METHACRYLATE	1573
METHYL LACTATE	1575
2-METHOXY PROPANOL-1	1576
m-CHLORONITROBENZENE	1582
2-HYDROXYACETOPHENONE	1583
4-HYDROXYACETOPHENONE	1584
p-PHENETIDINE	1587
ACETOL	1589
o-CHLOROPHENOL	1591
m-CHLOROPHENOL	1592
p-CHLOROPHENOL	1593
PENTAFLUOROPHENOL	1594
1-ISOPROPOXY-2-PROPANOL	1595
2-CHLOROETHANOL	1596
ISOPHTHALOYL CHLORIDE	1597
OCTAMETHYLTRISILOXANE	1630
OCTADECAMETHYLOCTASILOXANE	1631
TETRAETHYL SILANE	1632
HEXADECAMETHYLHEPTASILOXANE	1633
DIAMYL KETONE	1637
DIACETYL	1638
1-CYCLOHEXYLETHANONE	1640
3-METHYLCYCLOPENTANONE	1642
THYMOL	1649
1,2,3,6-TETRAHYDROBENZYL ALCOHOL	1652
7-METHYL-1-OCTENE	1656
CYCLOPROPANE CARBOXYLIC ACID	1668
ISOBUTYRIC ANHYDRIDE	1672
2,2,4-TRIMETHYL-1,3-PENTANEDIOL DIISOBUTYRATE	1673
ETHYL TRIMETHYL ACETATE	1676
DI(2-ETHYLHEXYL)TEREPHTHALATE	1677
ALLYL ACRYLATE	1679
DIALLYL PHTHALATE	1680

METHYL STEARATE	1681
ETHYL PHENYL ACETATE	1684
n-PENTYL n-BUTYRATE	1689
ISOPENTYL BUTYRATE	1690
ALLYLIDENE DIACETATE	1691
METHYL CAPRYLATE	1697
ISOPROPYL PALMITATE	1699
ETHYL CAPRATE	1700
ETHYL MYRISTATE	1702
DIMETHYL GLUTARATE	1716
DIMETHYL SUCCINATE	1718
DIOCTYL ADIPATE	1719
DI(2-ETHYLHEXYL)ADIPATE	1720
DIMETHYLMALONATE	1722
1,3,5-TRIETHYLBENZENE	1723
1-BROMODODECANE	1724
1,4-BIS(3-AMINOPROPYL)PIPERAZINE	1733
CYCLOPROPYL CYANIDE	1735
1,5-PENTANEDIAMINE	1739
2,4-DIMETHYLANILINE	1741
3-(N,N-DIMETHYLAMINO) PROPYLAMINE	1744
1-n-PENTYLNAPHTHALENE	1747
8-METHYLQUINOLINE	1748
TRI-n-OCTYLAMINE	1749
2-BUTOXIME	1750
N-METHYLPIPERIDINE	1751
IMIDAZOLE	1752
DIAMYLAMINE	1753
TRIAMYLAMINE	1754
UNDECYLAMINE	1755
TETRYL	1758
DI-n-OCTYLAMINE	1759
DI-2-ETHYLHEXYLAMINE	1760
TETRAMETHYLETHYLENEDIAMINE	1761
1,3-PROPANEDIAMINE	1763
N-ETHYLANILINE	1768
DI-tert-BUTYL DISULFIDE	1770
DI-tert-BUTYL SULFIDE	1771
2-METHYLTHIACYCLOPENTANE	1772
2-ETHYLTHIOPHENE	1779
2,5-DIMETHYLTHIOPHENE	1781
ETHYL PROPYL DISULFIDE	1783
METHYL PHENYL SULFIDE	1784
DI-n-BUTYL SULFIDE	1785

ETHYL PROPYL SULFIDE	1790
DIISOPROPYL SULFIDE	1796
DI-n-BUTYL DISULFIDE	1802
ETHYLENE GLYCOL 2-ETHYLHEXYL ETHER	1803
DICHLOROACETIC ACID	1806
PROPYLENE GLYCOL MONOMETHYL ETHER	1807
DIPROPYLENE GLYCOL MONOMETHYL ETHER	1808
PROPYLENE GLYCOL 1-tert-BUTYL ETHER	1809
TRIPROPYLENE GLYCOL MONOMETHYL ETHER	1810
p-CHLOROANILINE	1813
TRIETHYLENE GLYCOL n-BUTYL ETHER	1814
TRIETHYLENE GLYCOL MONOETHYL ETHER	1815
2,3-EPOXY-1-PROPANOL	1817
2-HEXOXYETHANOL	1818
2-(2-HEXOXYETHOXY)ETHANOL	1819
MONOOLEIN	1820
DIPROPYLENE GLYCOL n-PROPYL ETHER	1821
PROPYLENE GLYCOL n-PROPYL ETHER	1822
DIETHYLENE GLYCOL MONOBUTYL ETHER ACETATE	1823
PHENYLACETONITRILE	1827
N,N-DIETHYLHYDROXYLAMINE	1828
DIETHYLSULFITE	1836
1-ETHOXY-2-PROPANOL	1846
METHANESULFONYL CHLORIDE	1850
tert-BUTYL METHACRYLATE	1851
gamma-AMINOPROPYLTRIETHOXYSILANE	1860
METHYL DICHLOROSILANE	1866
METHYL TRICHLOROSILANE	1867
VINYLTRICHLOROSILANE	1868
TRIS(2-METHOXYETHOXY)VINYLSILANE	1870
METHYL SILICATE	1871
3-CHLOROPROPYLTRIMETHOXYSILANE	1872
[3-(2,3- EPOXYPROXY)PROPYL]TRIMETHOXYSILANE	1877
METHYLTRIACETOXYSILANE	1879
TRIETHYL ESTER PHOSPHOROUS ACID	1883
DIMETHYLCHLOROSILANE	1903
TRIMETHYLCHLOROSILANE	1904
DIMETHYLDICHLOROSILANE	1905
DECAMETHYLCYCLOPENTASILOXANE	1906
ETHYLTRICHLOROSILANE	1909
DICHLORODIETHYLSILANE	1910
DIGLYCOLIC ACID	1917
LEVULINIC ACID	1918
GUAIACOL	1920

ETHYLENE GLYCOL MONOPROPYL ETHER	1921
2-(2-BUTOXYETHOXY)ETHANOL	1922
4-CHLORO-3-NITROBENZOTRIFLUORIDE	1924
3,4-DICHLOROPHENYL ISOCYANATE	1925
3-NITROBENZOTRIFLUORIDE	1928
p-TOLUENESULFONIC ACID	1929
CHLOROACETALDEHYDE	1932
DICHLOROACETALDEHYDE	1933
1-CHLORO-2,4-DINITROBENZENE	1934
p-DIMETHYLAMINOBENZALDEHYDE	1935
ETHYL CHLOROFORMATE	1936
METHYL CHLOROFORMATE	1939
3,4-DICHLOROANILINE	1941
1,2-DICHLORO-4-NITROBENZENE	1942
DEXTROSE	1943
o-CHLORONITROBENZENE	1944
p-CHLORONITROBENZENE	1945
TRIETHYL PHOSPHATE	1946
CYCLOHEXANONE OXIME	1949
TRIMETHOXYSILANE	1952
DIMETHYLDIMETHOXYSILANE	1953
PHENYLTRICHLOROSILANE	1954
PHENYLMETHYLDICHLOROSILANE	1955
TETRADECAMETHYLHEXASILOXANE	1958
DECAMETHYLTETRASILOXANE	1959
PYRUVIC ACID	1960
TRI-o-CRESYL PHOSPHATE	1961
TRIPHENYL PHOSPHATE	1962
ETHYL ALUMINUM SESQUICHLORIDE	1963
3-AMINO-1-PROPANOL	1968
1-AMINO-2-PROPANOL	1969
METHYL CHLOROACETATE	1970
ETHYL THIOLACETATE	1971
ACETOACETANILIDE	1972
OXAZOLE	1973
ACETANILIDE	1974
LACTONITRILE	1976
DIMETHYL SULFATE	1978
DIETHYL SULFATE	1979
TARTARIC ACID	1985
ETHYL LACTATE	1987
2-ETHOXYETHYL ACETATE	1988
DIETHYLENE GLYCOL ETHYL ETHER ACETATE	1989
METHYL ACETOACETATE	1990

ETHYL ACETOACETATE	1991
METHYL CYANOACETATE	1992
ETHYL CYANOACETATE	1993
3-METHOXYPROPIONITRILE	1994
ISOXAZOLE	1995
TEREPHTHALOYL CHLORIDE	1998
1-CHLORO-2-PROPANOL	2001
PROPANESULFONYL CHLORIDE	2003
ETHANESULFONYL CHLORIDE	2004
tert-BUTYLFORMAMIDE	2005
N-METHYLACETAMIDE	2007
THIODIGLYCOL	2008
DIMERCAPTOETHYL ETHER	2010
2-MERCAPTOETHANOL	2011
ETHYLTHIOETHANOL	2012
1,2-ETHANEDITHIOL	2013
DIETHYLETHANOLAMINE	2014
METHYLETHANOLAMINE	2015
DIMETHYLETHANOLAMINE	2016
DIISOPROPANOLAMINE	2017
N-(2-HYDROXYETHYL)PIPERAZINE	2019
3-METHOXYISOPROPYLAMINE	2020
ETHYLENE GLYCOL MONOBUTYL ETHER ACETATE	2021
2-HYDROXYETHYL ACRYLATE	2023
2-HYDROXYPROPYL METHACRYLATE	2024
ETHYL-3-ETHOXYPROPIONATE	2025
2-HYDROXYPROPYL ACRYLATE	2027
DI-(2-CHLOROETHOXY)METHANE	2028
1-METHYLVINYL ACETATE	2030
2-PENTOXYETHANOL	2031
1-CHLORO-3-PROPANOL	2032
METHOXYACETONE	2034
METHYL para-TOLUATE	2041
PHENYL ACETATE	2042
PROPYLENE GLYCOL n-BUTYL ETHER	2044
DIPROPYLENE GLYCOL n-BUTYL ETHER	2045
TRIMETHYL BORATE	2049
t-BUTYL ACETOACETATE	2050
4-(2-AMINOETHYL)MORPHOLINE	2056
N-ETHYLMORPHOLINE	2057
METHYL-3-MERCAPTOPROPIONATE	2060
ETHYL LEVULINATE	2064
4-(2-HYDROXYETHYL)MORPHOLINE	2067
ACETOIN	2071

1,3-DIMETHYL-2-IMIDAZOLIDINONE	2073
N-ETHYL-2-PYRROLIDONE	2074
TRI-n-BUTYL PHOSPHATE	2077
5-(HYDROXYMETHYL) FURFURAL	2078
2-AMINO-2-METHYL-1-PROPANOL	2082
METHYL alpha-HYDROXYISOBUTYRATE	2091
4-METHYLGUAIACOL	2106
THIOACETIC-ACID	2118
3-BROMO-1-PROPENE	2120
3-IODO-1-PROPENE	2121
1-BROMOPENTANE	2133
2-CHLORO-2-METHYLBUTANE	2135

97

98 **Table S3.** The compounds correlated in the QSPR modeling for AIT

Name	Index
METHANE	0
ETHANE	1
PROPANE	2
ISOBUTANE	3
n-BUTANE	4
n-PENTANE	6
ISOPENTANE	7
NEOPENTANE	8
n-HEXANE	9
2-METHYLPENTANE	10
3-METHYLPENTANE	11
2,2-DIMETHYLBUTANE	12
2,3-DIMETHYLBUTANE	13
n-HEPTANE	14
2-METHYLHEXANE	15
3-METHYLHEXANE	16
2,3-DIMETHYLPENTANE	19
2,2,3-TRIMETHYLBUTANE	22
n-OCTANE	23
2-METHYL-3-ETHYLPENTANE	34
2,2,3-TRIMETHYLPENTANE	36
2,2,4-TRIMETHYLPENTANE	37
2,3,3-TRIMETHYLPENTANE	38
n-NONANE	41
3,3-DIETHYLPENTANE	45
2,2,3,3-TETRAMETHYLPENTANE	46
2,3,3,4-TETRAMETHYLPENTANE	49
n-DECANE	51

n-DODECANE	57
n-TETRADECANE	59
n-HEXADECANE	61
2-METHYLNONANE	77
2-METHYLOCTANE	81
3-METHYLOCTANE	82
4-METHYLOCTANE	83
CYCLOPROPANE	90
CYCLOPENTANE	92
METHYLCYCLOPENTANE	93
ETHYLCYCLOPENTANE	94
n-PROPYLCYCLOPENTANE	100
n-BUTYLCYCLOPENTANE	107
CYCLOHEXANE	115
METHYLCYCLOHEXANE	116
ETHYLCYCLOHEXANE	117
cis-1,2-DIMETHYLCYCLOHEXANE	119
trans-1,2-DIMETHYLCYCLOHEXANE	120
cis-1,3-DIMETHYLCYCLOHEXANE	121
trans-1,3-DIMETHYLCYCLOHEXANE	122
cis-1,4-DIMETHYLCYCLOHEXANE	123
trans-1,4-DIMETHYLCYCLOHEXANE	124
1-trans-3,5-TRIMETHYLCYCLOHEXANE	125
n-PROPYLCYCLOHEXANE	126
ISOPROPYLCYCLOHEXANE	127
n-BUTYLCYCLOHEXANE	129
cis-DECAHYDRONAPHTHALENE	130
trans-DECAHYDRONAPHTHALENE	131
BICYCLOHEXYL	132
2,4-DIMETHYL-3-ETHYLPENTANE	145
ETHYLENE	146
PROPYLENE	147
1-BUTENE	149
cis-2-BUTENE	150
trans-2-BUTENE	151
ISOBUTENE	152
1-PENTENE	154
cis-2-PENTENE	155
trans-2-PENTENE	156
2-METHYL-1-BUTENE	157
2-METHYL-2-BUTENE	159
1-HEXENE	161
cis-2-HEXENE	162
trans-2-HEXENE	163

2-METHYL-1-PENTENE	166
4-METHYL-1-PENTENE	168
2-ETHYL-1-BUTENE	174
2,3-DIMETHYL-1-BUTENE	175
2,3-DIMETHYL-2-BUTENE	177
1-HEPTENE	179
2,3,3-TRIMETHYL-1-BUTENE	193
1-OCTENE	195
2,4,4-TRIMETHYL-1-PENTENE	197
2,4,4-TRIMETHYL-2-PENTENE	198
1-DECENE	201
1-DODECENE	203
1-TETRADECENE	205
1-HEXADECENE	207
1-OCTADECENE	208
CYCLOPENTENE	210
CYCLOHEXENE	211
VINYLCYCLOHEXENE	224
d-LIMONENE	229
1,3-BUTADIENE	238
ISOPRENE	244
METHYLCYCLOPENTADIENE	247
CYCLOPENTADIENE	250
DICYCLOPENTADIENE	251
3-METHYL-1,4-PENTADIENE	256
ACETYLENE	267
METHYLACETYLENE	268
BENZENE	286
TOLUENE	287
ETHYLBENZENE	289
o-XYLENE	290
m-XYLENE	291
p-XYLENE	292
n-PROPYLBENZENE	294
CUMENE	295
o-ETHYLTOLUENE	296
m-ETHYLTOLUENE	297
p-ETHYLTOLUENE	298
1,2,3-TRIMETHYLBENZENE	299
1,2,4-TRIMETHYLBENZENE	300
MESITYLENE	301
n-BUTYLBENZENE	303
ISOBUTYLBENZENE	304
sec-BUTYLBENZENE	305

tert-BUTYLBENZENE	306
p-CYMENE	309
o-DIETHYLBENZENE	310
m-DIETHYLBENZENE	311
p-DIETHYLBENZENE	312
m-DIISOPROPYLBENZENE	327
p-DIISOPROPYLBENZENE	328
BIPHENYL	338
p-TERPHENYL	339
m-TERPHENYL	340
o-TERPHENYL	341
1,1-DIPHENYLETHANE	342
DIPHENYLMETHANE	343
1,2-DIPHENYLETHANE	344
STYRENE	378
p-METHYLSTYRENE	385
alpha-METHYLSTYRENE	386
NAPHTHALENE	395
1-METHYLNAPHTHALENE	396
1-ETHYLNAPHTHALENE	398
1,2,3,4-TETRAHYDRONAPHTHALENE	399
ANTHRACENE	424
2-NORBORNENE	441
alpha-PINENE	456
FORMALDEHYDE	474
ACETALDEHYDE	475
PROPANAL	476
BUTANAL	478
2-METHYLPROPANAL	479
PENTANAL	480
HEPTANAL	481
HEXANAL	482
OCTANAL	483
NONANAL	484
2-ETHYLHEXANAL	486
ACROLEIN	499
trans-CROTONALDEHYDE	501
BENZALDEHYDE	506
PARALDEHYDE	513
ACETONE	514
METHYL ETHYL KETONE	515
3-PENTANONE	516
METHYL ISOBUTYL KETONE	517
2-PENTANONE	523

2-HEXANONE	525
2-HEPTANONE	526
5-METHYL-2-HEXANONE	527
MESITYL OXIDE	528
DIISOBUTYL KETONE	530
N-METHYL-2-PYRROLIDONE	533
ACETYLACETONE	538
ISOPHORONE	539
CYCLOHEXANONE	541
ACETOPHENONE	544
epsilon-CAPROLACTONE	547
DIKETENE	553
METHANOL	555
ETHANOL	556
1-PROPANOL	557
ISOPROPANOL	558
1-BUTANOL	559
2-METHYL-1-PROPANOL	560
2-BUTANOL	561
2-METHYL-2-PROPANOL	562
1-PENTANOL	563
2-PENTANOL	564
2-METHYL-2-BUTANOL	565
2-METHYL-1-BUTANOL	566
1-HEXANOL	568
2-METHYL-1-PENTANOL	571
3-PENTANOL	574
2-ETHYL-1-HEXANOL	575
3-METHYL-1-BUTANOL	577
1-DECANOL	589
1-DODECANOL	592
CYCLOHEXANOL	603
1-METHYLCYCLOHEXANOL	604
cis-2-METHYLCYCLOHEXANOL	605
trans-2-METHYLCYCLOHEXANOL	606
cis-4-METHYLCYCLOHEXANOL	609
trans-4-METHYLCYCLOHEXANOL	610
TETRAHYDROFURFURYL ALCOHOL	618
ALLYL ALCOHOL	619
PROPARGYL ALCOHOL	631
BENZYL ALCOHOL	632
PHENOL	633
o-CRESOL	634
m-CRESOL	635

ALLYL ACETATE	750
ISOPROPYL ACETATE	751
VINYL ACETATE	753
METHYL PROPIONATE	754
ETHYL PROPIONATE	755
n-BUTYL PROPIONATE	757
ETHYL n-BUTYRATE	761
METHYL ACRYLATE	764
ETHYL ACRYLATE	765
n-BUTYL ACRYLATE	767
DIOCTYL PHTHALATE	774
n-PENTYL ACETATE	777
BENZYL ACETATE	779
ISOBUTYL ISOBUTYRATE	780
BENZYL BENZOATE	784
n-BUTYL BENZOATE	785
DIISODECYL PHTHALATE	791
METHYL SALICYLATE	793
DIISOBUTYL PHTHALATE	796
DIMETHYL TEREPHTHALATE	801
n-BUTYL STEARATE	803
DIBUTYL SEBACATE	804
2-ETHYLHEXYL ACRYLATE	806
ETHYLENE GLYCOL DIACETATE	807
ISOBUTYL METHACRYLATE	808
n-BUTYL METHACRYLATE	809
METHYL BENZOATE	810
ETHYL BENZOATE	811
DIMETHYL ETHER	821
DIETHYL ETHER	822
DIISOPROPYL ETHER	823
DI-n-BUTYL ETHER	824
METHYL tert-BUTYL ETHER	825
METHYL ETHYL ETHER	827
DI-n-HEXYL ETHER	832
DIVINYL ETHER	834
1,4-DIOXANE	839
TRIOXANE	840
DI-n-OCTYL ETHER	842
DI-n-PENTYL ETHER	843
METHYL n-PENTYL ETHER	847
METHYLAL	849
ACETAL	850
ETHYLENE OXIDE	859

1,2-PROPYLENE OXIDE	860
ETHYL VINYL ETHER	862
DI-n-PROPYL ETHER	863
BUTYL VINYL ETHER	864
TRIETHYLENE GLYCOL DIMETHYL ETHER	871
1,2-DIMETHOXYETHANE	872
DIETHYLENE GLYCOL DIMETHYL ETHER	873
DIETHYLENE GLYCOL DIETHYL ETHER	875
DIETHYLENE GLYCOL DI-n-BUTYL ETHER	876
DIPHENYL ETHER	882
METHYL VINYL ETHER	887
1,2-EPOXYBUTANE	888
CUMENE HYDROPEROXIDE	889
TETRAHYDROFURAN	896
BENZOYL PEROXIDE	898
METHYL CHLORIDE	912
ETHYL CHLORIDE	913
VINYL CHLORIDE	914
DICHLOROMETHANE	921
1,1-DICHLOROETHANE	924
1,2-DICHLOROETHANE	925
1,1,2-TRICHLOROETHANE	926
1,2-DICHLOROPROPANE	928
1,1,1-TRICHLOROETHANE	929
ISOPROPYL CHLORIDE	932
1,2,3-TRICHLOROPROPANE	934
TRICHLOROETHYLENE	939
3-CHLOROPROPENE	941
HEXACHLORO-1,3-BUTADIENE	943
BENZYL CHLORIDE	944
MONOCHLOROBENZENE	947
o-DICHLOROBENZENE	948
cis-1,2-DICHLOROETHYLENE	956
trans-1,2-DICHLOROETHYLENE	957
CHLOROPRENE	959
PROPYL CHLORIDE	960
n-BUTYL CHLORIDE	961
1-CHLOROPENTANE	963
1,1-DICHLOROETHYLENE	966
1,2,4-TRICHLOROBENZENE	967
CHLORODIFLUOROMETHANE	979
1,1-DIFLUOROETHYLENE	1000
TETRAFLUOROETHYLENE	1001
BROMOMETHANE	1007

BROMOETHANE	1008
1,1,2,2-TETRABROMOETHANE	1010
1-BROMOPROPANE	1011
1-BROMOBUTANE	1015
BROMOBENZENE	1021
DICHLOROFLUOROMETHANE	1031
METHYLAMINE	1036
DIMETHYLAMINE	1037
TRIMETHYLAMINE	1038
ETHYLAMINE	1039
TRIETHYLAMINE	1041
DI-n-PROPYLAMINE	1042
DIETHYLAMINE	1045
n-PROPYLAMINE	1046
n-BUTYLAMINE	1047
ISOBUTYLAMINE	1049
ISOPROPYLAMINE	1053
DIETHANOLAMINE	1058
CYCLOHEXYLAMINE	1062
DEHYDROABIETYLAMINE	1063
o-TOLUIDINE	1069
m-TOLUIDINE	1070
p-TOLUIDINE	1071
TRIETHYLENETETRAMINE	1072
ALLYLAMINE	1073
ETHYLENEDIAMINE	1074
ETHYLENEIMINE	1075
DIISOPROPYLAMINE	1076
QUINOLINE	1081
N,N-DIETHYLANILINE	1086
N-METHYLPYRROLE	1087
DIPHENYLAMINE	1089
PHENYLHYDRAZINE	1090
NITROMETHANE	1093
NITROETHANE	1094
1-NITROPROPANE	1095
2-NITROPROPANE	1096
HYDRACRYLONITRILE	1097
MORPHOLINE	1098
HYDROGEN CYANIDE	1103
ACETONITRILE	1104
ACRYLONITRILE	1106
ADIPONITRILE	1109
o-NITROTOLUENE	1110

p-NITROTOLUENE	1111
BUTYRONITRILE	1114
PYRIDINE	1121
ANILINE	1122
HEXAMETHYLENEIMINE	1124
N,N-DIMETHYLANILINE	1126
2-METHYLPYRIDINE	1127
ETHYL MERCAPTAN	1131
DIMETHYL SULFIDE	1149
DIMETHYL SULFOXIDE	1173
ACETYL CHLORIDE	1179
BENZOYL CHLORIDE	1184
DI(2-CHLOROETHYL)ETHER	1194
3-(METHYLMERCAPTO)PROPANAL	1196
N,N-DIMETHYLFORMAMIDE	1204
epsilon-CAPROLACTAM	1208
alpha-EPICHLOROHYDRIN	1209
ACETONE CYANOHYDRIN	1210
NITROBENZENE	1214
ACETALDOL	1216
FURFURAL	1217
PROPYLENE GLYCOL MONOMETHYL ETHER ACETATE	1224
CARBON DISULFIDE	1252
HEXAMETHYLDISILOXANE	1263
DODECAMETHYLPENTASILOXANE	1267
METHYL VINYL DICHLOROSILANE	1269
TETRAMETHYLSILANE	1278
OCTAMETHYLCYCLOTETRASILOXANE	1281
2,3-DIMETHYLOCTANE	1302
alpha-METHYLBENZYL ALCOHOL	1316
1,4-CYCLOHEXANEDIMETHANOL	1328
2-ETHYL HEXANOIC ACID	1339
n-HEPTANOIC ACID	1340
FUMARIC ACID	1347
2-ETHYL BUTYRIC ACID	1355
CYCLOHEXYL ACETATE	1377
GLYCERYL TRIACETATE	1379
DIETHYL PHTHALATE	1382
DI-n-BUTYL PHTHALATE	1383
DIMETHYL PHTHALATE	1384
ISOBUTYL ACRYLATE	1390
DIETHYL MALEATE	1392
1,3-DIOXANE	1403
1,2-DIETHOXYETHANE	1405

2,3-DIHYDROFURAN	1406
VINYL BROMIDE	1462
1-CHLORO-1,1-DIFLUOROETHANE	1463
VINYL FLUORIDE	1464
n-OCTYLAMINE	1467
DIETHYLENETRIAMINE	1475
TETRAETHYLENEPENTAMINE	1476
CYCLOPENTYLAMINE	1487
N-AMINOETHYL ETHANOLAMINE	1490
PIPERAZINE	1510
NITROGLYCERINE	1519
o-NITROANILINE	1520
p-AMINODIPHENYL	1527
METHYL ISOCYANATE	1532
3-METHYLPYRIDINE	1536
4-METHYLPYRIDINE	1538
CAMPHOR	1550
DIACETONE ALCOHOL	1554
FURFURYL ALCOHOL	1555
N,N-DIMETHYLACETAMIDE	1556
p-METHOXYPHENOL	1559
2-METHOXYETHANOL	1560
2-ETHOXYETHANOL	1561
2-BUTOXYETHANOL	1562
2-(2-METHOXYETHOXY)ETHANOL	1563
2-(2-ETHOXYETHOXY)ETHANOL	1564
METHYL LACTATE	1575
2-CHLOROETHANOL	1596
4-FORMYLMORPHOLINE	1598
OCTAMETHYLTRISILOXANE	1630
TETRAETHYL SILANE	1632
2,3-EPOXY-1-PROPANOL	1817
DIETHYLENE GLYCOL MONOBUTYL ETHER ACETATE	1823
METHYL DICHLOROSILANE	1866
VINYLTRICHLOROSILANE	1868
TETRAETHYL LEAD	1889
METHYL SILANE	1900
DIMETHYL SILANE	1901
TRIMETHYL SILANE	1902
TRIMETHYLCHLOROSILANE	1904
DECAMETHYLCYCLOPENTASILOXANE	1906
DIGLYCOLIC ACID	1917
ETHYLENE GLYCOL MONOPROPYL ETHER	1921
2-(2-BUTOXYETHOXY)ETHANOL	1922

1-CHLORO-2,4-DINITROBENZENE	1934
ETHYL CHLOROFORMATE	1936
METHYL CHLOROFORMATE	1939
3,4-DICHLOROANILINE	1941
TRIETHYL PHOSPHATE	1946
DIMETHYLDIMETHOXYSILANE	1953
TETRADECAMETHYLHEXASILOXANE	1958
DECAMETHYLTETRASILOXANE	1959
TRI-o-CRESYL PHOSPHATE	1961
1-AMINO-2-PROPANOL	1969
ACETOACETANILIDE	1972
ACETANILIDE	1974
DIMETHYL SULFATE	1978
DIETHYL SULFATE	1979
CITRIC ACID	1983
TARTARIC ACID	1985
ETHYL LACTATE	1987
2-ETHOXYETHYL ACETATE	1988
DIETHYLENE GLYCOL ETHYL ETHER ACETATE	1989
METHYL ACETOACETATE	1990
ETHYL ACETOACETATE	1991
ETHYL HYDROGEN SULFATE	1997
2-MERCAPTOETHANOL	2011
METHYLETHANOLAMINE	2015
DIMETHYLETHANOLAMINE	2016
DIISOPROPANOLAMINE	2017
ETHYLENE GLYCOL MONOBUTYL ETHER ACETATE	2021
ETHYL-3-ETHOXYPROPIONATE	2025
DL-2,3-butanediol	2274
2-ethylbutanoic acid	2275
tetrakis(trimethylsilyloxy)silane	2276
2-methylbiphenyl	2277
tributyrin	2278
butyl citrate	2279
4-octylphenyl salicylate	2280

Table S4. The compounds correlated in the QSPR modeling for LFL

Name	Index
METHANE	0
PROPANE	2
ISOBUTANE	3
n-BUTANE	4
n-PENTANE	6

ISOPENTANE	7
NEOPENTANE	8
n-HEXANE	9
2-METHYLPENTANE	10
3-METHYLPENTANE	11
2,2-DIMETHYLBUTANE	12
2,3-DIMETHYLBUTANE	13
n-HEPTANE	14
2-METHYLHEXANE	15
2,3-DIMETHYLPENTANE	19
2,2,3-TRIMETHYLBUTANE	22
n-OCTANE	23
2,2,4-TRIMETHYLPENTANE	37
n-NONANE	41
3,3-DIETHYLPENTANE	45
2,2,3,3-TETRAMETHYLPENTANE	46
n-DECANE	51
n-DODECANE	57
n-TRIDECANE	58
n-TETRADECANE	59
CYCLOPROPANE	90
CYCLOBUTANE	91
ETHYLCYCLOPENTANE	94
n-PROPYLCYCLOPENTANE	100
CYCLOHEXANE	115
METHYLCYCLOHEXANE	116
ETHYLCYCLOHEXANE	117
cis-DECAHYDRONAPHTHALENE	130
trans-DECAHYDRONAPHTHALENE	131
BICYCLOHEXYL	132
1,1-DIETHYLCYCLOHEXANE	133
CYCLOHEPTANE	135
PROPYLENE	147
1-BUTENE	149
cis-2-BUTENE	150
trans-2-BUTENE	151
ISOBUTENE	152
1-PENTENE	154
2-METHYL-1-BUTENE	157
2-METHYL-2-BUTENE	159
1-HEXENE	161
1-HEPTENE	179
trans-2-HEPTENE	181
1-OCTENE	195

trans-2-OCTENE	196
2,4,4-TRIMETHYL-1-PENTENE	197
1-DECENE	201
1-DODECENE	203
CYCLOHEXENE	211
VINYLCYCLOHEXENE	224
d-LIMONENE	229
1,3-BUTADIENE	238
ISOPRENE	244
METHYLCYCLOPENTADIENE	247
1,4-HEXADIENE	248
DICYCLOPENTADIENE	251
METHYLACETYLENE	268
DIMETHYLACETYLENE	270
1-HEXYNE	276
VINYLACETYLENE	279
BENZENE	286
TOLUENE	287
ETHYLBENZENE	289
o-XYLENE	290
m-XYLENE	291
p-XYLENE	292
n-PROPYLBENZENE	294
CUMENE	295
1,2,3-TRIMETHYLBENZENE	299
MESITYLENE	301
n-BUTYLBENZENE	303
ISOBUTYLBENZENE	304
sec-BUTYLBENZENE	305
tert-BUTYLBENZENE	306
p-CYMENE	309
p-DIETHYLBENZENE	312
BIPHENYL	338
STYRENE	378
m-METHYLSTYRENE	380
5-o-TOLYL-2-PENTENE	381
p-METHYLSTYRENE	385
alpha-METHYLSTYRENE	386
m-DIVINYLBENZENE	387
NAPHTHALENE	395
1-METHYLNAPHTHALENE	396
1,2,3,4-TETRAHYDRONAPHTHALENE	399
ANTHRACENE	424
VINYLNORBORNENE	431

sec-BUTYLCYCLOHEXANE	445
alpha-PINENE	456
ACETALDEHYDE	475
PROPANAL	476
BUTANAL	478
2-METHYLPROPANAL	479
2-ETHYLHEXANAL	486
ACROLEIN	499
cis-CROTONALDEHYDE	500
trans-CROTONALDEHYDE	501
METHACROLEIN	502
BENZALDEHYDE	506
PARALDEHYDE	513
ACETONE	514
METHYL ETHYL KETONE	515
3-PENTANONE	516
METHYL ISOBUTYL KETONE	517
2-PENTANONE	523
METHYL ISOPROPYL KETONE	524
2-HEPTANONE	526
5-METHYL-2-HEXANONE	527
MESITYL OXIDE	528
DIISOBUTYL KETONE	530
N-METHYL-2-PYRROLIDONE	533
2-NONANONE	536
CYCLOHEXANONE	541
ACETOPHENONE	544
beta-PROPIOLACTONE	545
gamma-BUTYROLACTONE	546
METHYL ISOPROPENYL KETONE	550
DIKETENE	553
METHANOL	555
ETHANOL	556
1-PROPANOL	557
ISOPROPANOL	558
1-BUTANOL	559
2-METHYL-1-PROPANOL	560
2-BUTANOL	561
2-METHYL-2-PROPANOL	562
1-PENTANOL	563
2-PENTANOL	564
2-METHYL-1-BUTANOL	566
1-HEXANOL	568
2-METHYL-1-PENTANOL	571

3-PENTANOL	574
2-ETHYL-1-HEXANOL	575
3-METHYL-1-BUTANOL	577
3-METHYL-2-BUTANOL	578
4-METHYL-2-PENTANOL	583
1-OCTANOL	585
1-NONANOL	587
1-DECANOL	589
2-ETHYL-1-BUTANOL	599
CYCLOHEXANOL	603
1-METHYLCYCLOHEXANOL	604
alpha-TERPINEOL	614
TETRAHYDROFURFURYL ALCOHOL	618
ALLYL ALCOHOL	619
2,6-XYLENOL	628
PROPARGYL ALCOHOL	631
o-CRESOL	634
m-CRESOL	635
p-CRESOL	636
ETHYLENE GLYCOL	643
DIETHYLENE GLYCOL	644
TRIETHYLENE GLYCOL	645
1,2-PROPYLENE GLYCOL	653
1,3-PROPYLENE GLYCOL	654
DIPROPYLENE GLYCOL	655
NEOPENTYL GLYCOL	656
2-BUTYNE-1,4-DIOL	657
2-METHYL-1,3-PROPANEDIOL	660
1,3-BUTANEDIOL	662
HEXYLENE GLYCOL	663
GLYCEROL	665
cis-2-BUTENE-1,4-DIOL	673
trans-2-BUTENE-1,4-DIOL	674
1,4-BUTANEDIOL	675
1,3-BENZENEDIOL	679
FORMIC ACID	685
ACETIC ACID	686
PROPIONIC ACID	687
n-BUTYRIC ACID	690
n-PENTANOIC ACID	692
ISOBUTYRIC ACID	694
ISOVALERIC ACID	695
trans-CROTONIC ACID	708
ACRYLIC ACID	711

METHACRYLIC ACID	712
ADIPIC ACID	719
ACETIC ANHYDRIDE	725
PROPIONIC ANHYDRIDE	726
BUTYRIC ANHYDRIDE	727
PHTHALIC ANHYDRIDE	731
METHYL FORMATE	734
ETHYL FORMATE	735
n-PROPYL FORMATE	736
n-BUTYL FORMATE	737
ISOBUTYL FORMATE	738
n-OCTYL FORMATE	740
METHYL ACETATE	744
n-PROPYL ACETATE	746
n-BUTYL ACETATE	747
ISOBUTYL ACETATE	748
ISOPENTYL ACETATE	749
ALLYL ACETATE	750
VINYL ACETATE	753
METHYL PROPIONATE	754
ETHYL PROPIONATE	755
METHYL ACRYLATE	764
ETHYL ACRYLATE	765
n-BUTYL ACRYLATE	767
METHYL METHACRYLATE	771
ETHYL METHACRYLATE	772
DIOCTYL PHTHALATE	774
DIISOOCTYL PHTHALATE	775
n-PENTYL ACETATE	777
2-ETHYLHEXYL ACETATE	778
ISOBUTYL ISOBUTYRATE	780
n-HEXYL ACETATE	783
BENZYL BENZOATE	784
ETHYLENE CARBONATE	786
n-DECYL ACETATE	790
DIISODECYL PHTHALATE	791
DIISOBUTYL PHTHALATE	796
2-ETHYLHEXYL ACRYLATE	806
METHYL BENZOATE	810
ETHYL BENZOATE	811
DIETHYL CARBONATE	812
DIMETHYL ETHER	821
DIETHYL ETHER	822
DIISOPROPYL ETHER	823

DI-n-BUTYL ETHER	824
METHYL tert-BUTYL ETHER	825
METHYL ETHYL ETHER	827
DIVINYL ETHER	834
ETHYL PROPYL ETHER	835
1,4-DIOXANE	839
TRIOXANE	840
tert-BUTYL ETHYL ETHER	846
METHYLAL	849
ACETAL	850
1,2-EPOXY-2-METHYLPROPANE	858
ETHYLENE OXIDE	859
1,2-PROPYLENE OXIDE	860
1,3-PROPYLENE OXIDE	861
ETHYL VINYL ETHER	862
1,2-DIMETHOXYETHANE	872
ANISOLE	878
PHENETOLE	879
DIPHENYL ETHER	882
METHYL VINYL ETHER	887
FURAN	895
TETRAHYDROFURAN	896
DI-t-BUTYL PEROXIDE	899
METHYL CHLORIDE	912
ETHYL CHLORIDE	913
VINYL CHLORIDE	914
1,4-DICHLORO-trans-2-BUTENE	915
DICHLOROMETHANE	921
1,1-DICHLOROETHANE	924
1,2-DICHLOROETHANE	925
1,1,2-TRICHLOROETHANE	926
1,2-DICHLOROPROPANE	928
1,1,1-TRICHLOROETHANE	929
ISOPROPYL CHLORIDE	932
ISOBUTYL CHLORIDE	935
TRICHLOROETHYLENE	939
3-CHLOROPROPENE	941
BENZYL CHLORIDE	944
MONOCHLOROBENZENE	947
o-DICHLOROBENZENE	948
o-CHLOROTOLUENE	953
p-CHLOROTOLUENE	954
cis-1,2-DICHLOROETHYLENE	956
trans-1,2-DICHLOROETHYLENE	957

CHLOROPRENE	959
PROPYL CHLORIDE	960
n-BUTYL CHLORIDE	961
1-CHLOROPENTANE	963
1,1-DICHLOROETHYLENE	966
1,2,4-TRICHLOROBENZENE	967
2-CHLOROPROPENE	970
3,4-DICHLORO-1-BUTENE	972
CHLOROFLUOROMETHANE	975
PENTAFLUOROETHYL METHYL ETHER	983
METHYL FLUORIDE	986
DIFLUOROMETHANE	987
ETHYL FLUORIDE	990
1,1,1-TRIFLUOROETHANE	991
1,1-DIFLUOROETHYLENE	1000
TETRAFLUOROETHYLENE	1001
1,1-DIFLUOROETHANE	1006
BROMOMETHANE	1007
BROMOETHANE	1008
2-BROMOPROPANE	1012
1-BROMOBUTANE	1015
METHYL IODIDE	1022
METHYLAMINE	1036
DIMETHYLAMINE	1037
TRIMETHYLAMINE	1038
ETHYLAMINE	1039
TRIETHYLAMINE	1041
DI-n-PROPYLAMINE	1042
DIETHYLAMINE	1045
n-PROPYLAMINE	1046
n-BUTYLAMINE	1047
n-PENTYLAMINE	1048
HYDRAZINE	1051
ISOPROPYLAMINE	1053
PYRROLE	1055
DIETHANOLAMINE	1058
tert-BUTYLAMINE	1061
CYCLOHEXYLAMINE	1062
HEXAMETHYLENEDIAMINE	1064
ALLYLAMINE	1073
ETHYLENEDIAMINE	1074
ETHYLENEIMINE	1075
DI-n-BUTYLAMINE	1077
NITROMETHANE	1093

NITROETHANE	1094
1-NITROPROPANE	1095
2-NITROPROPANE	1096
HYDRACRYLONITRILE	1097
MORPHOLINE	1098
PYRROLIDINE	1099
ACETONITRILE	1104
PROPIONITRILE	1105
ACRYLONITRILE	1106
METHACRYLONITRILE	1107
o-NITROTOLUENE	1110
p-NITROTOLUENE	1111
BUTYRONITRILE	1114
VALERONITRILE	1115
HEXANENITRILE	1117
ISOBUTYRONITRILE	1118
trans-CROTONITRILE	1119
BENZONITRILE	1120
PYRIDINE	1121
ANILINE	1122
2,4-TOLUENE DIISOCYANATE	1123
HEXAMETHYLENEIMINE	1124
N,N-DIMETHYLANILINE	1126
2-METHYLPYRIDINE	1127
cis-CROTONITRILE	1128
CYANOGEN	1129
METHYL MERCAPTAN	1130
ETHYL MERCAPTAN	1131
DIMETHYL SULFIDE	1149
THIOPHENE	1150
TETRAHYDROTHIOPHENE	1172
DIMETHYL SULFOXIDE	1173
ACETYL CHLORIDE	1179
BENZOYL CHLORIDE	1184
DI(2-CHLOROETHYL)ETHER	1194
N,N-DIMETHYLFORMAMIDE	1204
alpha-EPICHLOROHYDRIN	1209
ACETONE CYANOHYDRIN	1210
FURFURAL	1217
PROPYLENE GLYCOL MONOMETHYL ETHER ACETATE	1224
HEXAMETHYLDISILOXANE	1263
HEXAMETHYLCYCLOTRISILOXANE	1264
TETRAMETHYLSILANE	1278
OCTAMETHYLCYCLOTETRASILOXANE	1281

2,6-DIMETHYL-4-HEPTANOL	1310
2-ETHYL HEXANOIC ACID	1339
n-HEXYL FORMATE	1365
CYCLOHEXYL ACETATE	1377
PROPYLENE CARBONATE	1378
DIETHYL PHTHALATE	1382
DI-n-BUTYL PHTHALATE	1383
DIMETHYL PHTHALATE	1384
ISOBUTYL ACRYLATE	1390
DIISOBUTYL ETHER	1399
1,1-DICHLOROPROPANE	1408
2,3-DICHLOROPROPENE	1410
3,3,3-TRIFLUOROPROPENE	1419
1,1,2-TRIFLUOROETHANE	1420
1,1,2,2,3-PENTAFLUOROPROPANE	1426
1,2-DIFLUOROETHANE	1436
1,1-DICHLORO-1-FLUOROETHANE	1441
p-BROMOTOLUENE	1452
CHLOROTRIFLUOROETHYLENE	1459
VINYL BROMIDE	1462
1-CHLORO-1,1-DIFLUOROETHANE	1463
VINYL FLUORIDE	1464
n-OCTYLAMINE	1467
n-NONYLAMINE	1468
TRI-n-BUTYLAMINE	1474
TRIPROPYLAMINE	1477
n-BUTYL ISOCYANATE	1480
o-PHENYLENEDIAMINE	1483
DIALLYLAMINE	1497
p-PHENYLENEDIAMINE	1508
3-METHYLPYRIDINE	1536
METHYLGLUTARONITRILE	1537
4-METHYLPYRIDINE	1538
1,6-HEXAMETHYLENE DIISOCYANATE	1539
CAMPHOR	1550
N-METHYLFORMAMIDE	1552
DIACETONE ALCOHOL	1554
FURFURYL ALCOHOL	1555
N,N-DIMETHYLACETAMIDE	1556
2-METHOXYETHANOL	1560
2-ETHOXYETHANOL	1561
2-BUTOXYETHANOL	1562
2-(2-METHOXYETHOXY)ETHANOL	1563
2-(2-ETHOXYETHOXY)ETHANOL	1564

2-AMINOETHOXYETHANOL	1565
METHYL LACTATE	1575
2-CHLOROETHANOL	1596
OCTAMETHYLTRISILOXANE	1630
ISOBUTYRIC ANHYDRIDE	1672
3-(N,N-DIMETHYLAMINO) PROPYLAMINE	1744
TETRAMETHYLETHYLENEDIAMINE	1761
BENZIDINE	1766
N-ETHYLANILINE	1768
PROPYLENE GLYCOL 1-tert-BUTYL ETHER	1809
DIETHYLENE GLYCOL MONOPROPYL ETHER	1812
p-CHLOROANILINE	1813
DIPROPYLENE GLYCOL n-PROPYL ETHER	1821
DIETHYLENE GLYCOL MONOBUTYL ETHER ACETATE	1823
1-ETHOXY-2-PROPANOL	1846
TETRAETHOXYSILANE	1859
METHYL TRICHLOROSILANE	1867
VINYLTRICHLOROSILANE	1868
TRIMETHYLCHLOROSILANE	1904
DIMETHYLDICHLOROSILANE	1905
DICHLORODIETHYLSILANE	1910
ETHYLENE GLYCOL MONOPROPYL ETHER	1921
2-(2-BUTOXYETHOXY)ETHANOL	1922
1-CHLORO-2,4-DINITROBENZENE	1934
METHYL CHLOROFORMATE	1939
3,4-DICHLOROANILINE	1941
TRIETHYL PHOSPHATE	1946
CYCLOHEXANONE OXIME	1949
VINYLTRIMETHOXYSILANE	1951
DIMETHYLDIMETHOXYSILANE	1953
TRIMETHYL SILANOL	1957
1-AMINO-2-PROPANOL	1969
METHYL CHLOROACETATE	1970
DIMETHYL SULFATE	1978
ETHYL LACTATE	1987
2-ETHOXYETHYL ACETATE	1988
DIETHYLENE GLYCOL ETHYL ETHER ACETATE	1989
ETHYL ACETOACETATE	1991
3-METHOXYPROPIONITRILE	1994
2-MERCAPTOETHANOL	2011
ETHYLENE GLYCOL MONOBUTYL ETHER ACETATE	2021
ETHYL-3-ETHOXYPROPIONATE	2025
1-METHYLVINYL ACETATE	2030

Table S5. The compounds correlated in the QSPR modeling for UFL

Name	Index
METHANE	0
PROPANE	2
n-BUTANE	4
n-PENTANE	6
ISOPENTANE	7
NEOPENTANE	8
2-METHYLPENTANE	10
3-METHYLPENTANE	11
2,2-DIMETHYLBUTANE	12
2,3-DIMETHYLBUTANE	13
n-HEPTANE	14
2-METHYLHEXANE	15
2,2-DIMETHYLPENTANE	18
2,3-DIMETHYLPENTANE	19
n-OCTANE	23
2,2,4-TRIMETHYLPENTANE	37
n-NONANE	41
3,3-DIETHYLPENTANE	45
2,2,3,3-TETRAMETHYLPENTANE	46
n-DECANE	51
CYCLOBUTANE	91
METHYLCYCLOPENTANE	93
ETHYLCYCLOPENTANE	94
CYCLOHEXANE	115
METHYLCYCLOHEXANE	116
ETHYLCYCLOHEXANE	117
cis-DECAHYDRONAPHTHALENE	130
trans-DECAHYDRONAPHTHALENE	131
BICYCLOHEXYL	132
1,1-DIETHYLCYCLOHEXANE	133
CYCLOHEPTANE	135
PROPYLENE	147
1-BUTENE	149
cis-2-BUTENE	150
trans-2-BUTENE	151
ISOBUTENE	152
1-PENTENE	154
2-METHYL-2-BUTENE	159
1-OCTENE	195
2,4,4-TRIMETHYL-1-PENTENE	197
1-DECENE	201
1-DODECENE	203

CYCLOHEXENE	211
d-LIMONENE	229
1,3-BUTADIENE	238
ISOPRENE	244
METHYLCYCLOPENTADIENE	247
1,4-HEXADIENE	248
DICYCLOPENTADIENE	251
VINYLACETYLENE	279
BENZENE	286
TOLUENE	287
ETHYLBENZENE	289
o-XYLENE	290
m-XYLENE	291
p-XYLENE	292
CUMENE	295
1,2,3-TRIMETHYLBENZENE	299
1,2,4-TRIMETHYLBENZENE	300
MESITYLENE	301
n-BUTYLBENZENE	303
ISOBUTYLBENZENE	304
sec-BUTYLBENZENE	305
tert-BUTYLBENZENE	306
p-CYMENE	309
p-DIETHYLBENZENE	312
BIPHENYL	338
2-PHENYLBUTENE-1	362
STYRENE	378
m-METHYLSTYRENE	380
p-METHYLSTYRENE	385
alpha-METHYLSTYRENE	386
m-DIVINYLBENZENE	387
NAPHTHALENE	395
1-METHYLNAPHTHALENE	396
1,2,3,4-TETRAHYDRONAPHTHALENE	399
sec-BUTYLCYCLOHEXANE	445
ACETALDEHYDE	475
PROPANAL	476
BUTANAL	478
2-METHYLPROPANAL	479
2-ETHYLHEXANAL	486
ACROLEIN	499
trans-CROTONALDEHYDE	501
BENZALDEHYDE	506
ACETONE	514

METHYL ETHYL KETONE	515
3-PENTANONE	516
METHYL ISOBUTYL KETONE	517
3-HEXANONE	522
2-PENTANONE	523
METHYL ISOPROPYL KETONE	524
2-HEXANONE	525
2-HEPTANONE	526
5-METHYL-2-HEXANONE	527
MESITYL OXIDE	528
DIISOBUTYL KETONE	530
N-METHYL-2-PYRROLIDONE	533
2-NONANONE	536
CYCLOPENTANONE	540
CYCLOHEXANONE	541
gamma-BUTYROLACTONE	546
METHYL ISOPROPENYL KETONE	550
ETHANOL	556
1-PROPANOL	557
ISOPROPANOL	558
1-BUTANOL	559
2-METHYL-1-PROPANOL	560
2-BUTANOL	561
2-METHYL-2-PROPANOL	562
1-PENTANOL	563
2-PENTANOL	564
2-METHYL-1-BUTANOL	566
2-METHYL-1-PENTANOL	571
3-PENTANOL	574
2-ETHYL-1-HEXANOL	575
3-METHYL-1-BUTANOL	577
4-METHYL-2-PENTANOL	583
1-NONANOL	587
CYCLOHEXANOL	603
TETRAHYDROFURFURYL ALCOHOL	618
ALLYL ALCOHOL	619
1,2-PROPYLENE GLYCOL	653
NEOPENTYL GLYCOL	656
2-BUTYNE-1,4-DIOL	657
HEXYLENE GLYCOL	663
GLYCEROL	665
cis-2-BUTENE-1,4-DIOL	673
trans-2-BUTENE-1,4-DIOL	674
1,4-BUTANEDIOL	675

METHYL ETHYL ETHER	827
METHYL n-PROPYL ETHER	828
DIVINYL ETHER	834
ETHYL PROPYL ETHER	835
1,4-DIOXANE	839
METHYL tert-PENTYL ETHER	845
tert-BUTYL ETHYL ETHER	846
METHYLAL	849
1,2-EPOXY-2-METHYLPROPANE	858
1,2-PROPYLENE OXIDE	860
1,3-PROPYLENE OXIDE	861
1,2-DIMETHOXYETHANE	872
DIPHENYL ETHER	882
1,2-EPOXYBUTANE	888
CUMENE HYDROPEROXIDE	889
FURAN	895
TETRAHYDROFURAN	896
METHYL CHLORIDE	912
ETHYL CHLORIDE	913
VINYL CHLORIDE	914
DICHLOROMETHANE	921
1,2-DICHLOROETHANE	925
1,1,2-TRICHLOROETHANE	926
1,2-DICHLOROPROPANE	928
1,1,1-TRICHLOROETHANE	929
ISOPROPYL CHLORIDE	932
ISOBUTYL CHLORIDE	935
3-CHLOROPROPENE	941
o-DICHLOROBENZENE	948
o-CHLOROTOLUENE	953
p-CHLOROTOLUENE	954
cis-1,2-DICHLOROETHYLENE	956
trans-1,2-DICHLOROETHYLENE	957
CHLOROPRENE	959
PROPYL CHLORIDE	960
n-BUTYL CHLORIDE	961
1-CHLOROPENTANE	963
1,1-DICHLOROETHYLENE	966
1,2,4-TRICHLOROBENZENE	967
2-CHLOROPROPENE	970
3,4-DICHLORO-1-BUTENE	972
CHLOROFLUOROMETHANE	975
PENTAFLUOROETHYL METHYL ETHER	983
METHYL FLUORIDE	986

DIFLUOROMETHANE	987
ETHYL FLUORIDE	990
1,1,1-TRIFLUOROETHANE	991
1,1-DIFLUOROETHYLENE	1000
TETRAFLUOROETHYLENE	1001
1,1-DIFLUOROETHANE	1006
BROMOMETHANE	1007
BROMOETHANE	1008
1-BROMOPROPANE	1011
1-BROMOBUTANE	1015
DIMETHYLAMINE	1037
TRIMETHYLAMINE	1038
ETHYLAMINE	1039
TRIETHYLAMINE	1041
DI-n-PROPYLAMINE	1042
DIETHYLAMINE	1045
n-PROPYLAMINE	1046
n-BUTYLAMINE	1047
ISOBUTYLAMINE	1049
HYDRAZINE	1051
DIISOBUTYLAMINE	1052
ISOPROPYLAMINE	1053
PYRROLE	1055
MONOETHANOLAMINE	1057
tert-BUTYLAMINE	1061
HEXAMETHYLENEDIAMINE	1064
ALLYLAMINE	1073
ETHYLENEIMINE	1075
DIISOPROPYLAMINE	1076
DI-n-BUTYLAMINE	1077
QUINOLINE	1081
1,2-PROPANEDIAMINE	1085
NITROMETHANE	1093
2-NITROPROPANE	1096
HYDRACRYLONITRILE	1097
MORPHOLINE	1098
PYRROLIDINE	1099
ACETONITRILE	1104
PROPIONITRILE	1105
ACRYLONITRILE	1106
METHACRYLONITRILE	1107
ADIPONITRILE	1109
HEXANENITRILE	1117
ISOBUTYRONITRILE	1118

BENZONITRILE	1120
PYRIDINE	1121
ANILINE	1122
2,4-TOLUENE DIISOCYANATE	1123
HEXAMETHYLENEIMINE	1124
N,N-DIMETHYLANILINE	1126
2-METHYLPYRIDINE	1127
CYANOGEN	1129
METHYL MERCAPTAN	1130
ETHYL MERCAPTAN	1131
DIMETHYL SULFIDE	1149
THIOPHENE	1150
TETRAHYDROTHIOPHENE	1172
DIMETHYL SULFOXIDE	1173
ACETYL CHLORIDE	1179
BENZOYL CHLORIDE	1184
FLUOROBENZENE	1188
N,N-DIMETHYLFORMAMIDE	1204
alpha-EPICHLOROHYDRIN	1209
ACETONE CYANOHYDRIN	1210
FURFURAL	1217
PROPYLENE GLYCOL MONOMETHYL ETHER ACETATE	1224
HEXAMETHYLDISILOXANE	1263
HEXAMETHYLCYCLOTRISILOXANE	1264
TETRAMETHYLSILANE	1278
2,6-DIMETHYL-4-HEPTANOL	1310
2-ETHYL HEXANOIC ACID	1339
PROPYLENE CARBONATE	1378
2,3-DICHLOROPROPENE	1410
3,3,3-TRIFLUOROPROPENE	1419
1,1,2-TRIFLUOROETHANE	1420
1,1,2,2,3-PENTAFLUOROPROPANE	1426
1,2-DIFLUOROETHANE	1436
1,1-DICHLORO-1-FLUOROETHANE	1441
CHLOROTRIFLUOROETHYLENE	1459
VINYL BROMIDE	1462
1-CHLORO-1,1-DIFLUOROETHANE	1463
VINYL FLUORIDE	1464
n-OCTYLAMINE	1467
n-NONYLAMINE	1468
TRI-n-BUTYLAMINE	1474
TRIPROPYLAMINE	1477
n-BUTYL ISOCYANATE	1480
DIALLYLAMINE	1497

ISOQUINOLINE	1525
3-METHYLPYRIDINE	1536
METHYLGLUTARONITRILE	1537
4-METHYLPYRIDINE	1538
1,6-HEXAMETHYLENE DIISOCYANATE	1539
CAMPHOR	1550
N-METHYLFORMAMIDE	1552
DIACETONE ALCOHOL	1554
FURFURYL ALCOHOL	1555
N,N-DIMETHYLACETAMIDE	1556
2-METHOXYETHANOL	1560
2-ETHOXYETHANOL	1561
2-BUTOXYETHANOL	1562
2-(2-METHOXYETHOXY)ETHANOL	1563
2-(2-ETHOXYETHOXY)ETHANOL	1564
2-AMINOETHOXYETHANOL	1565
2-METHYL-, 3-HYDROXY-2,2,4- TRIMETHYLPENTYL PROPANOATE	1588
2-CHLOROETHANOL	1596
ISOBUTYRIC ANHYDRIDE	1672
3-(N,N-DIMETHYLAMINO) PROPYLAMINE	1744
TETRAMETHYLETHYLENEDIAMINE	1761
N-ETHYLANILINE	1768
PROPYLENE GLYCOL 1-tert-BUTYL ETHER	1809
2-(2-HEXOXYETHOXY)ETHANOL	1819
DIPROPYLENE GLYCOL n-PROPYL ETHER	1821
DIETHYLENE GLYCOL MONOBUTYL ETHER ACETATE	1823
1-ETHOXY-2-PROPANOL	1846
VINYLTRICHLOROSILANE	1868
DECAMETHYLCYCLOPENTASILOXANE	1906
ETHYLENE GLYCOL MONOPROPYL ETHER	1921
1-CHLORO-2,4-DINITROBENZENE	1934
METHYL CHLOROFORMATE	1939
3,4-DICHLOROANILINE	1941
TRIETHYL PHOSPHATE	1946
VINYLTRIMETHOXYSILANE	1951
PHENYLTRICHLOROSILANE	1954
3-AMINO-1-PROPANOL	1968
1-AMINO-2-PROPANOL	1969
METHYL CHLOROACETATE	1970
DIMETHYL SULFATE	1978
ETHYL LACTATE	1987
2-ETHOXYETHYL ACETATE	1988
DIETHYLENE GLYCOL ETHYL ETHER ACETATE	1989
ETHYL ACETOACETATE	1991

S5. Training the proposed MDNN

The joint training and alternative training are often employed in the modeling of multi-task deep learning network. The joint training requires that all the data of multiple target properties in one row are available, which is calculated according to the loss values and standard deviation of residuals of each task. The joint training is used to update all parameters simultaneously in the MDNN when the joint loss is back-propagated. The whole process is shown in **Figure S3**.

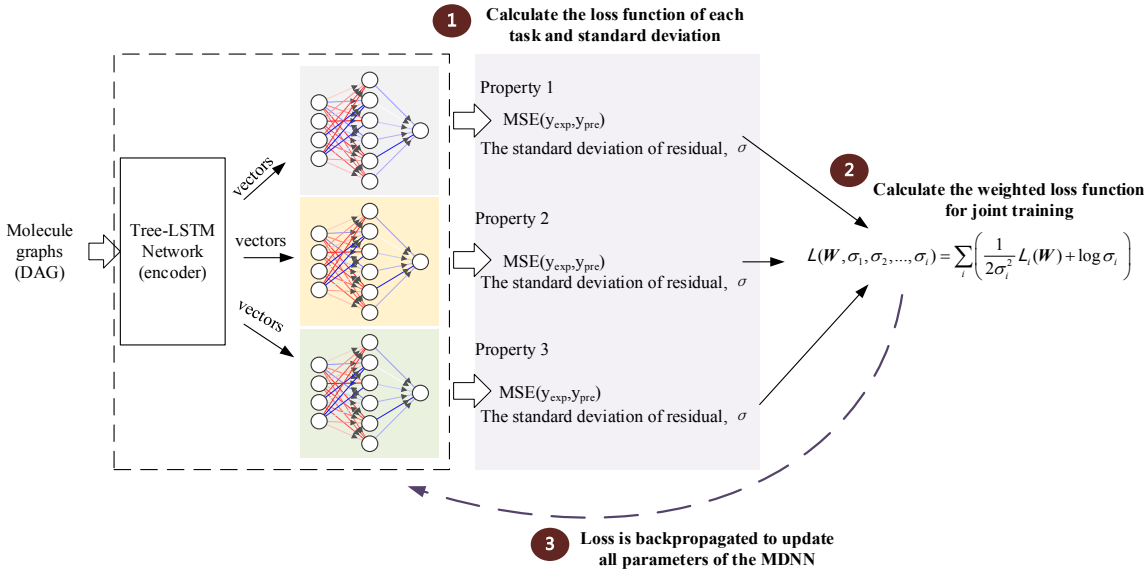

Figure S3. The flowchart of joint training in the proposed MDNN (the solid arrow means feedforward and the dashed arrow means backpropagation)

The alternative training was employed to train blocks of the MDNN in rotation. In this strategy of training, the loss gradients of each task are respectively back-propagated. The parameters of Tree-LSTM network are updated by the loss gradients with different weights of residual dispersion. And each FNN only used to the corresponding target property are trained based on the loss of each task without the weight. The whole flowchart of alternative training is presented in **Figure S4**. In this study, the joint training and alternative training were both employed to train the MDNN as shown in **Figure S5**.

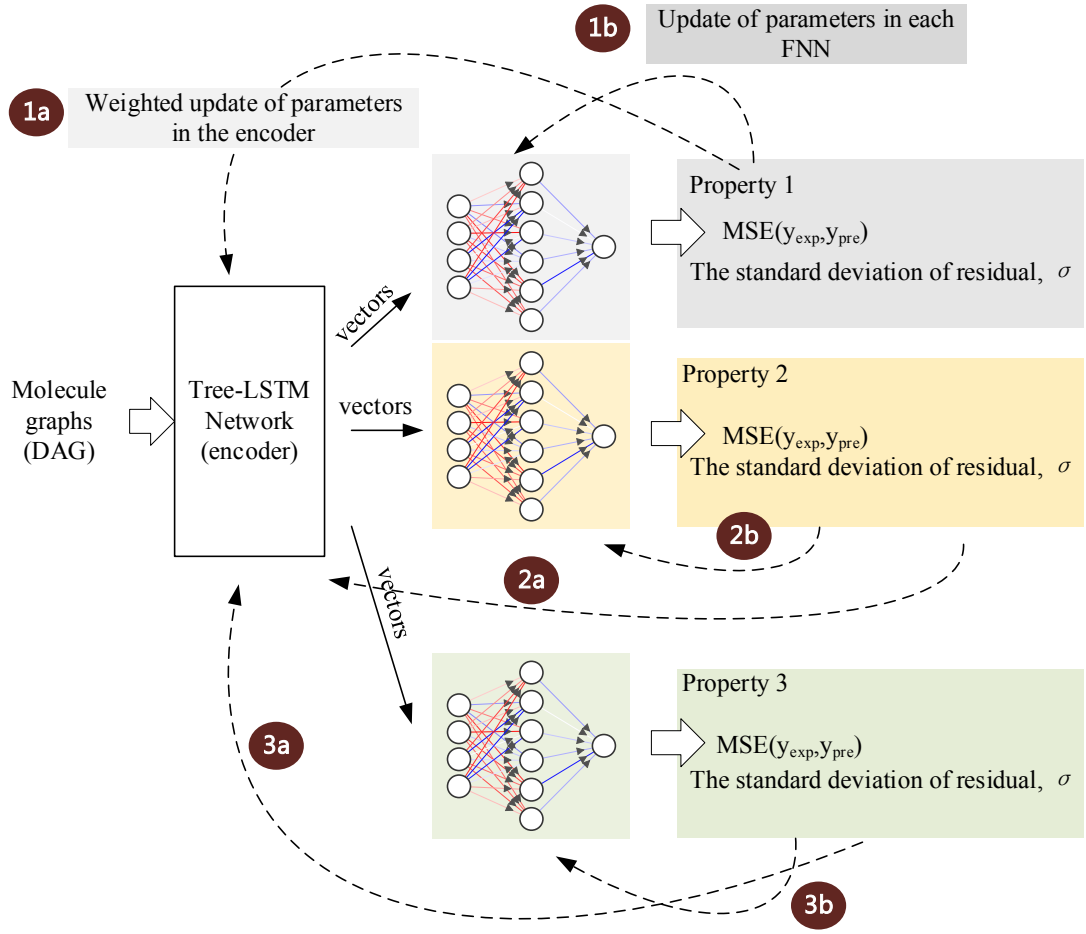

Figure S4. The flowchart of alternative training in the proposed MDNN (the solid arrow means feedforward and the dashed arrow means backpropagation)

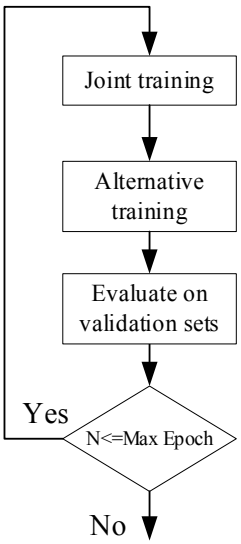

Figure S5. The schematic diagram of the training workflow combined the joint and alternative training

S6. Comparison between the proposed model and existing models

To investigate the performance of the MDNN model, the learned model with previous prediction models for the four properties was compared. Even though it is difficult to compare the learned MDNN model with the previous prediction models fully and fairly, because each model might be regressed from different data sets involving different set sizes, different categories of compounds, or different molecular descriptors.

Flash point temperature

The final correlation results obtained in this work presents an r of 0.9808 for the training set, an r of 0.9643 and an r of 0.9554 for the validation set and test set respectively. As compared with other previous studies, the values of *MPE* and R^2 indicates that the multi-task model could provide competitive ability in prediction (**Table S6**). The proposed MDNN model was trained on the training set including 822 pure compounds, and it also tested on the test set involving 177 pure compounds. The predicted results of the test set prove that the proposed model can estimate the FPT values for other compounds that not included in the training set, although the value of R^2 decreases slightly. It should be noted that the linear or non-linear models proposed by Frutiger *et al.*[7], Hukkerikar *et al.* [8], Suzuki[9] and Khajeh and Modarress [10] had not been validated on the external test sets. As a result, the transferability of NLR and MLR models is not verified on extra test sets, though they outperformed on the training sets. By contrast, the models based on classical machine learning (*i.e.*, ANN, ANFIS and SVM) proposed by Gharagheizi *et al.* [11], Pan *et al.*[12], Khajeh and Modarress[10], Albahri[13] and Patel *et al.*[14] also showed good accuracy, while the models of Patel *et al.*[14], Albahri[13] and Khajeh and Modarress[10] were obtained from less compounds. Notably, Gharagheizi *et al.* [11] obtained the model on the data set including both experimental and predicted values of FPT.

Table S6. The comparison of the proposed model with other existing models for predicting FPT.

Model	Model type	*MPE* (%)	R^2	r	No. of data
The proposed model in this work	MDNN	3.120	0.9541	0.9808	822 (train)
		3.401	0.9214	0.9643	177 (validation)
		4.001	0.9022	0.9554	177 (test)
Frutiger *et al.*[7]	MLR-GC	2.0	0.9900	-	927 (train)
Hukkerikar *et al.* [8]	MLR-GC	2.8	0.9671	-	512 (train)
Hukkerikar *et al.* [8]	MLR-GC+	3.27	0.7951	-	512 (train)

Suzuki[9]	NLR-QSPR	-	-	0.9670	400 (train)
Alibakhshi *et al*.[15]	MLR-GC	1.225	-	0.9935	740 (train)
Khajeh and Modarress[10]	ANFIS-QSPR	-	0.9690	-	95 (train)
Gharagheizi *et al*.[11]	ANN-GC	-	0.9767	-	1241 (train)
		-	0.9661	-	137 (test)
Pan *et al*.[12]	SVM-QSPR	-	0.9800	-	1026 (train)
		-	0.9510	-	256 (test)
Albahri[13]	ANN-GC	1.12	-	0.9961	335 (train)
		0.97	-	0.9980	40 (test)
Patel *et al*.[14]	ANN-QSPR	-	0.8980	-	189 (train)
		-	0.6620	-	46 (test)

*Note: the correlation coefficient *r* was applied in this study.

Seemingly, the value of R^2 or *r* obtained by the proposed MDNN model for the test set is lower than the one obtained in the studies involving Gharagheizi et al.[11], Pan *et al*.[12] and Albahri[13]. Of note is that the performance metrics of MDNN model were calculated with outliers. Albahri[13] obtained the better results on a smaller data set. In a word, the proposed MDNN model can provide competitive accuracy of FPT prediction with the three referenced models.

Auto-ignition temperature

For the property of AIT, the proposed MDNN model was trained on the training set of 232 compounds and validated on 49 compounds, and finally tested on an external test set of 54 compounds. The resulting correlation factor *r* for the training set is 0.9421 and for the validation set it is 0.9135, while the *r* is 0.8020 for the test set. The statistical indicators of previous models compared with the proposed MDNN are provided in **Table S7**, while the number of data points is also presented for the available test set. It can be found that the performance of the MDNN is at the same level as the existing models in terms of *MPE* and R^2. The MLR-based models proposed by Frutiger[7], Hukkerikar[8] and Suzuki[16] were not validated by an external test set in their studies. Among the previous studies based on ANN or SVM, the ANN-GC model of Gharagheizi[17] showed the best accuracy with the most samples involving both experimental and predicted values. Notably, Frutiger et al.[7] excluded the estimated values in the regression and only employed 513 compounds with experimental points. Both Albahri and George[18] and Pan et al.[17,19] also proposed similar strategies to correlate the property AIT but with fewer

175 substances.

176 **Table S7**. The comparison of the proposed model with other existing models for predicting AIT.

Models	Model types	*MPE*	R^2	*r*	No. of data
The proposed model in this work	MDNN	5.367	0.8845	0.9421	232 (train)
		6.443	0.8173	0.9135	49 (validation)
		9.234	0.5985	0.8020	54 (test)
Hukkerikar *et al.* [8]	MLR-GC	2.09	0.9742	-	570 (train)
Frutiger *et al.*[7]	MLR-GC	6.4	0.76	-	513 (train)
Suzuki[16]	MLR-QSPR	-	-	0.881	250 (train)
Albahri and George[18]	ANN-GC	2.6	0.98	-	470 (train)
		2.9	0.98	-	20 (test)
Gharagheizi[17]	ANN-GC	1.6	0.984	-	821 (train)
		1.6	0.986	-	102 (test)
Pan *et al.*[19]	SVM-QSPR	-	0.901	-	356 (train)
		-	0.874	-	90 (test)
Pan *et al.*[20]	ANN-QSPR	-	-	0.9741	76 (train)
		-	-	0.9063	42 (test)

177 ***Lower flammable limit***

178 As can be seen from the statistical indicators presented in **Table S8**, the *MPE* and R^2 values
179 of both training and test sets resulted in the proposed method are competitive with other
180 reported methods for the LFL property. Considering the obtained *MPE*, the MDNN model
181 performs better than the previous LFL models of Frutiger *et al.*[7] with the same data set.
182 Mendiburu *et al.*[24] reported a decent model only for C-H compounds at some specified
183 temperatures, however, their model cannot be compared directly with the MDNN model since
184 they employed far less samples on one family of chemicals. The methods of Albahri[25], Rowley
185 *et al.*[22], Seaton[23] and Lazzús[26] performed slightly better.

186 **Table S8**. The comparison of the proposed model with other existing models for predicting LFL.

Models	Model type	*MPE* (%)	R^2	No. of data
The current model	MDNN	16.27	0.9877	306 (train)
		19.87	0.9619	67 (validation)
		19.26	0.7435	67 (test)
Frutiger *et al.*[7]	MLR-GC	15.9	0.91	443 (train)
Rowley *et al.*[22]	MLR-GC	10.67	-	509 (train)
Seaton[23]	NLR-GC	10.0	-	152 (train)

Albahri[25]	MLR-GC	4.1	0.93	472 (train)
		0.02	-	18 (test)
Gharagheizi[27]	MLR-QSPR	6.70	0.9698	845 (train)
		6.65	0.9728	211 (test)
Gharagheizi[28]	ANN-GC	4.35	0.99	846 (train)
		5.70	0.971	211 (test)
Lazzus[26]	ANN-GC	8.6	0.9876	328 (train)
		8.5	0.9819	90 (test)
Mendiburu *et al*.[24]	NLR-QSPR	-	0.9652	60 (train)
		-	0.9239	60 (test)
Pan *et al*.[29]	SVM-QSPR	-	0.979	830 (train)
		-	0.979	208 (test)

Gharagheizi[27,28] and Pan *et al*.[29] correlated more compounds but involving experimental and prediction values. In fact, there are also many small values of LFL, which could be a reason of the high relative residuals.

Upper flammable limit

Among most published studies, the prediction accuracy of the existing models for UFLs is lower than the accuracy of the models for LFLs. The lower accuracy of these models for the UFLs could be related to the lager discreteness of the available data set. According to **Table S7**, the correlation and prediction of UFL achieved by the MDNN is compared with the previous methods. As with the models for other properties, the correlation models proposed by Frutiger *et al*. [7], Albahri[25], Gharagheizi[30], Pan *et al*.[31], Lazzús[26], High and Danner[32], Seaton[23] and Mendiburu[33] require the information related to molecular structures such as the occurrence of molecular groups, numerical molecular descriptors or other manually extracted features of molecular structures. Hereinto, the best correlation with R^2 of 0.9776 for the total set of flammable compounds were achieved by Lazzús[26] but with fewer substances employed than the compounds employed in the method reported by Gharagheizi[30]. However, Gharagheizi[30] employed the predicted and experimental values simultaneously.

Table S9. The comparison of the proposed model with the previous models for predicting UFL.

Models	Model types	*MPE*	R^2	No. of data
The current model	MDNN	18.77	0.9238	232 (train)
		18.99	0.5703	49 (validation)
		24.15	0.6868	54 (test)

Frutiger *et al.*[7]	MLR-GC	15.9	0.91	351 (train)
High and Danner[32]	NLR-GC	26.4	-	181 (train)
Seaton[23]	NLR-GC	20.0	-	152 (train)
Albahri[25]	MLR-GC	11.8	0.96	475 (train)
		0.41	-	13 (test)
Pan *et al.*[31]	GA-MLR-QSPR	-	0.758	579 (train)
Gharagheizi[30]	GA-MLR-QSPR	9.55	0.9202	693 (train)
		9.58	0.9199	172 (test)
Lazzus[26]	PSO-ANN-GC	7.0	0.9780	328 (train)
		7.5	0.9771	90 (test)
Mendiburu *et al.*[33]	combustion theory and chemical equilibrium	9.66	0.9382	136 (train)
		7.63	0.9188	72 (test)

S7. The applicability domains and outlier detection

The convex hulls define the interpolation regions in the target properties and molecular feature vectors. The boundary of a convex hull describes the smallest convex area covering the training and validation sets. It is easy to calculate the convex hull covering a date set in the software package, SciPy[37]. The results of calculation were visualized by a python implementation of VTK[38]. **Figures S6-S9** show the convex hulls representing the ADs of four properties in the initially learned model (Ⅰ).

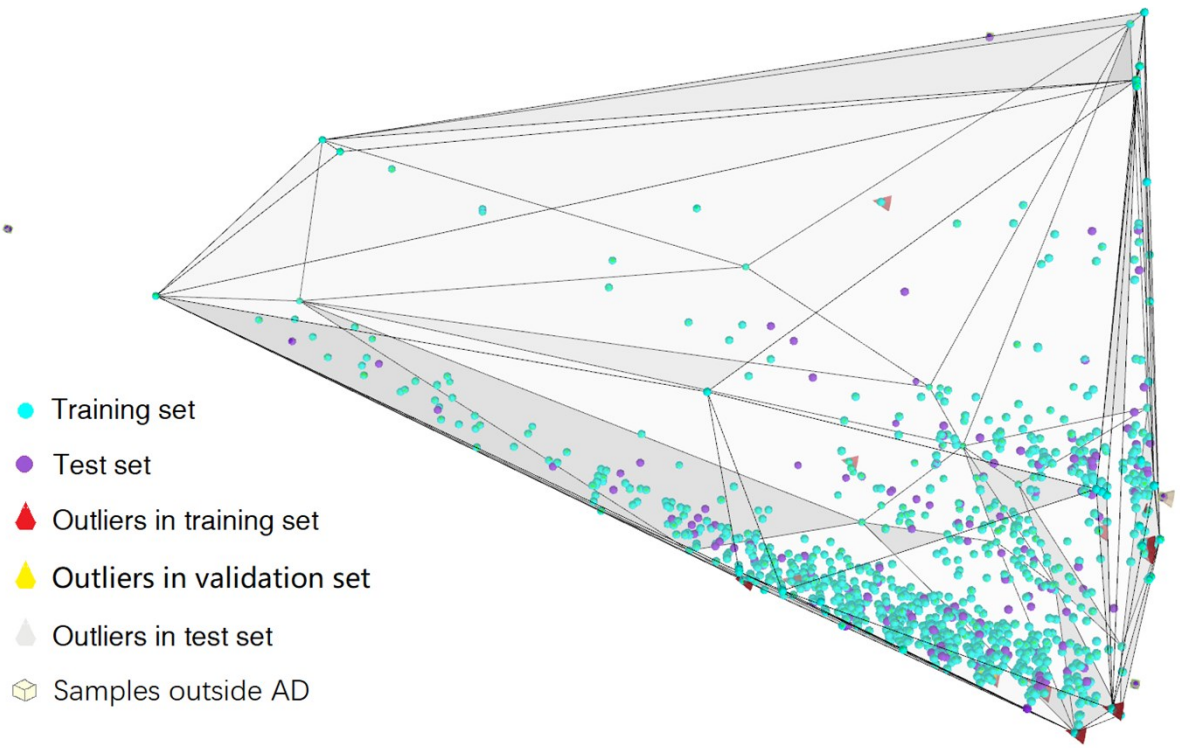

Figure S6. The convex hull representing the AD for FPT

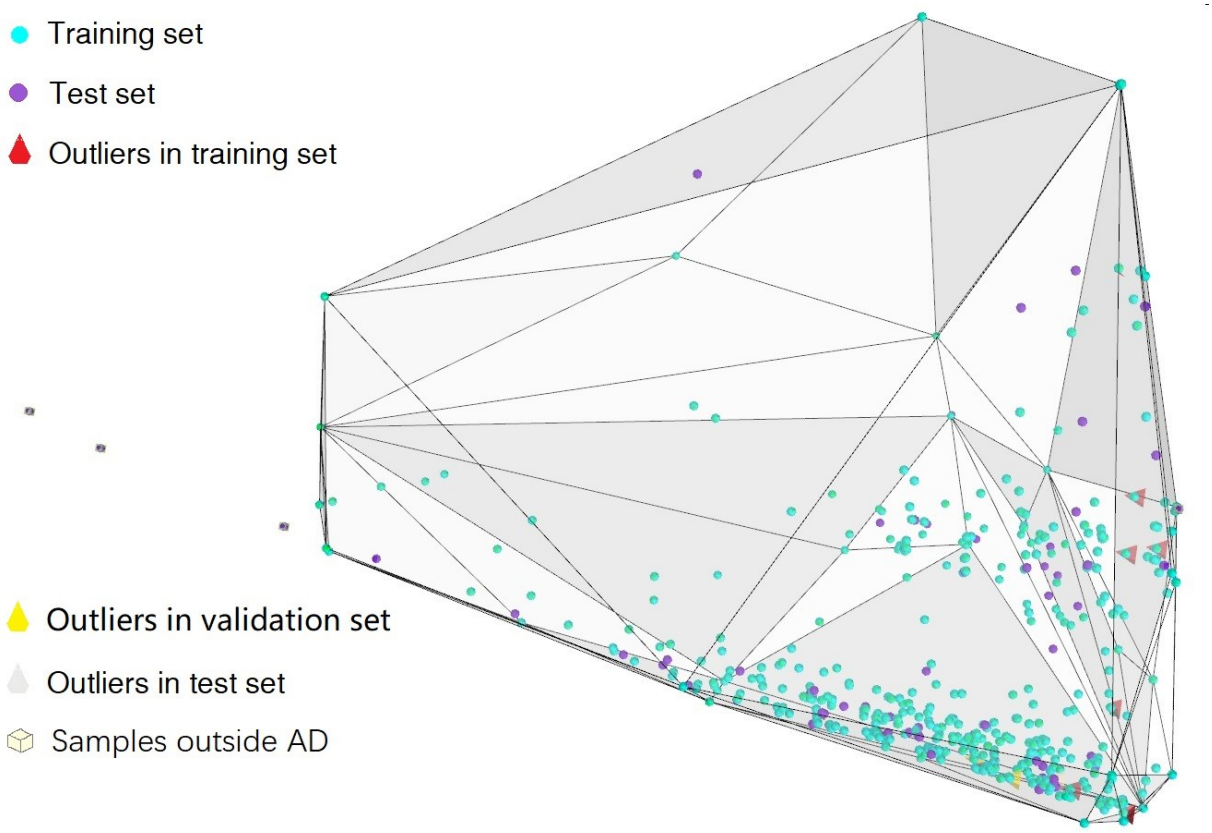

Figure S7. The convex hull representing the AD for AIT

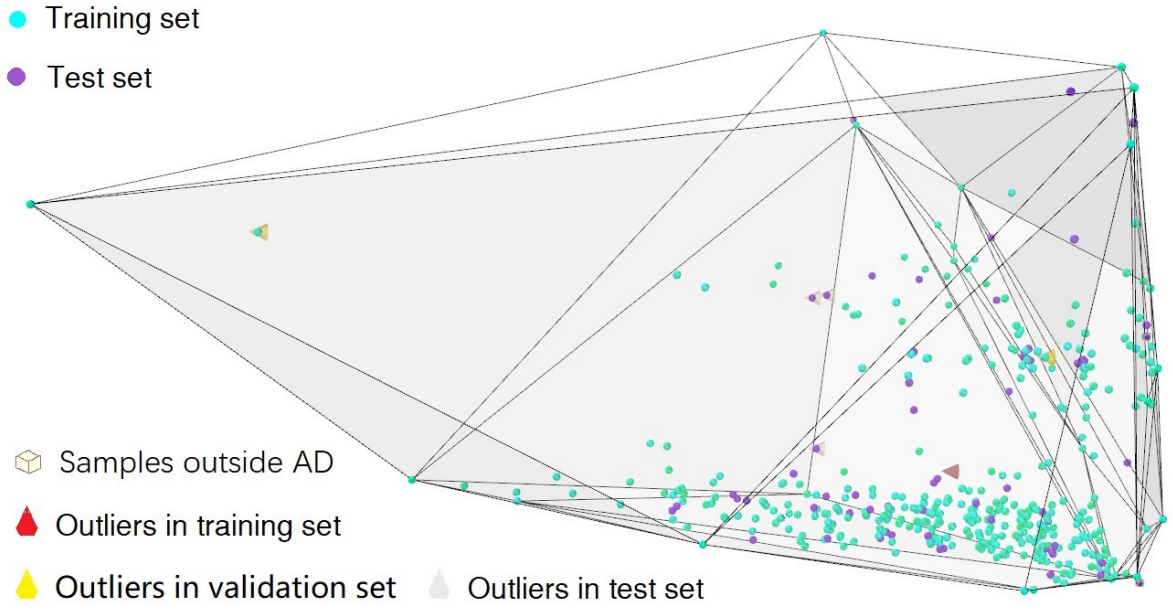

Figure S8. The convex hull representing the AD for LFL

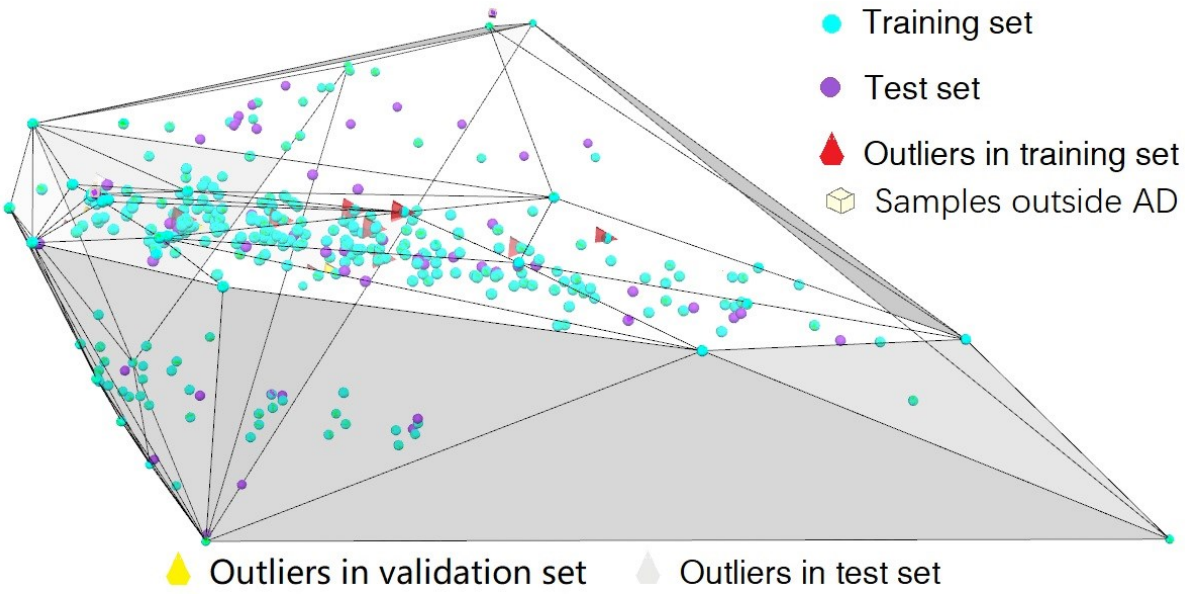

Figure S9. The convex hull representing the AD for UFL

The residual ECDFs of four properties are shown in **Figure S10** for the model (Ⅰ).

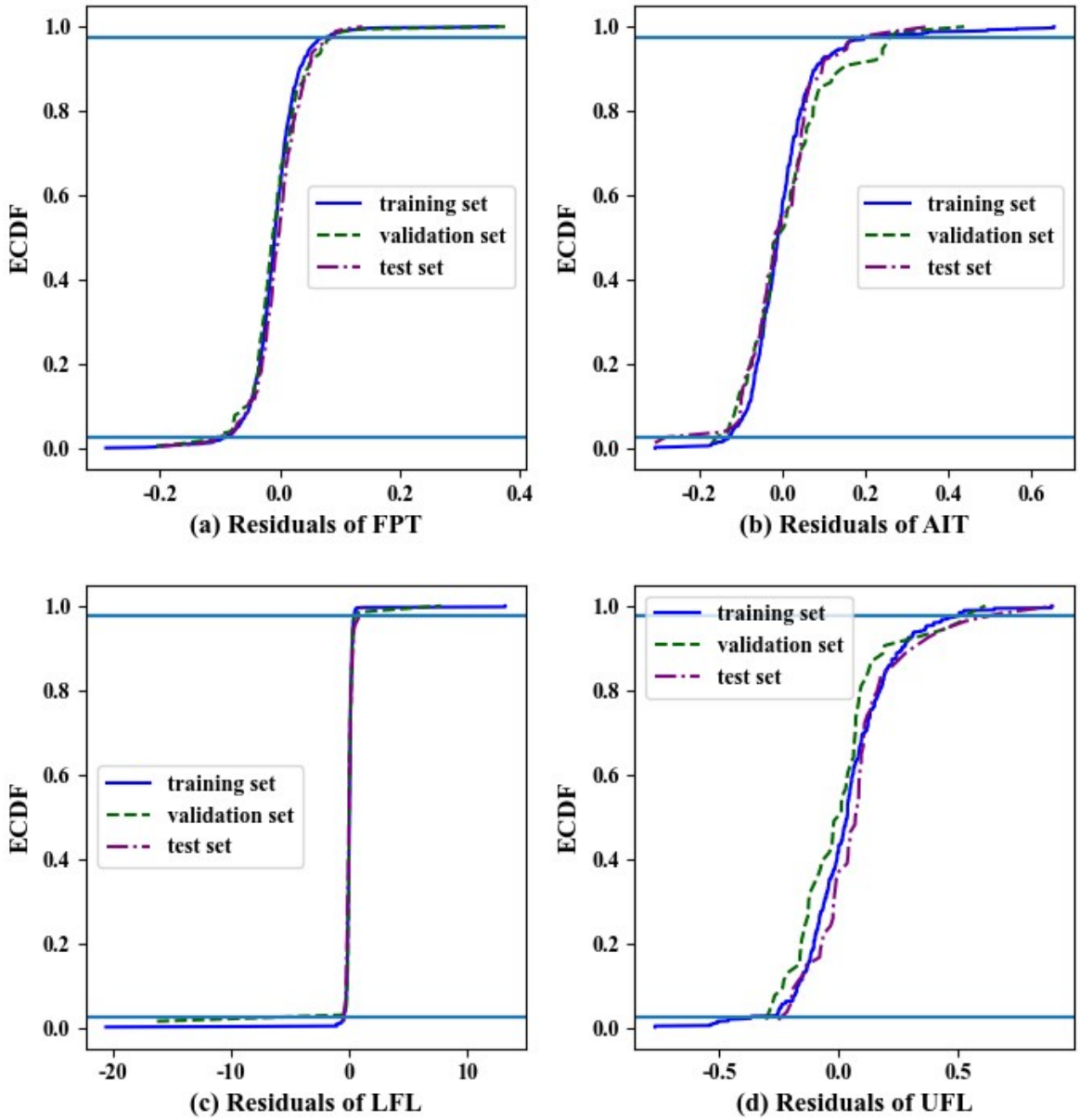

Figure S10. The plots of residual ECDFs for the training, validation, and test sets of: (a) FPT; (b) AIT; (c) LFL; (d) UFL.

The outlier compounds were identified by ECDF from the training and validation sets, which are listed in **Table S10**. As these compounds were excluded from the training and validation sets, the ADs were narrowed.

Table S10. The outliers identified from the training and validation sets.

FPT	AIT	LFL	UFL
ISOPROPYL IODIDE	DIETHANOLAMINE	DIISODECYL PHTHALATE	1,2-PROPYLENE OXIDE
N,N-	CITRIC ACID	DIOCTYL	1-ETHOXY-2-

DIMETHYLFORMAMIDE		PHTHALATE	PROPANOL
DIMETHYLCHLOROSILANE	FUMARIC ACID	METHYLGLUTARONITRILE	TETRAMETHYLSILANE
1,5,9-CYCLODODECATRIENE	FORMIC ACID	DIISOOCTYL PHTHALATE	1,3-PROPYLENE OXIDE
DIMETHYL SULFOXIDE	PYRENE	m-DIVINYLBENZENE	PHENYLTRICHLOROSILANE
DIVINYL ETHER	TRI-n-OCTYLAMINE	-	1,2-EPOXYBUTANE
tert-BUTYLFORMAMIDE	NICOTINONITRILE	-	NEOPENTYL GLYCOL
PROPYLENE CARBONATE	ETHANE	-	DIETHYLENE GLYCOL ETHYL ETHER ACETATE
1,2,3-TRIMETHYLINDENE	FORMAMIDE	-	TETRAMETHYLETHYLENEDIAMINE
n-PROPYL IODIDE	TRIAMYLAMINE	-	BROMOETHANE
TETRAETHYL SILANE	TRI-o-CRESYL PHOSPHATE	-	METHYLGLUTARONITRILE
N,N-DIMETHYLACETAMIDE	PROPIONIC ANHYDRIDE	-	MONOETHANOLAMINE
N-METHYLFORMAMIDE	5-METHYL-2-HEXANONE	-	2-(2-HEXOXYETHOXY)ETHANOL
ETHYLENE CARBONATE	2-(2-BUTOXYETHOXY)ETHANOL	-	DIPROPYLENE GLYCOL n-PROPYL ETHER
3-IODO-1-PROPENE	ISOPRENE	-	1,2-DIMETHOXYETHANE
N-METHYLACETAMIDE	METHYL ACETOACETATE	-	TRIMETHYLAMINE
IMIDAZOLE	TETRAFLUOROETHYLENE	-	BENZOYL CHLORIDE
HYDRAZINE	ACETALDEHYDE	-	1,1-DIFLUOROETHYLENE
FORMAMIDE	DIGLYCOLIC ACID	-	DECAMETHYLCYCLOPENTASILOXANE
3-BROMO-1-PROPENE	BENZALDEHYDE	-	2-BUTYNE-1,4-DIOL
ISOPHORONE DIISOCYANATE	o-NITROTOLUENE	-	VINYL BROMIDE

SULFOLANE	PHENYLHYDRAZINE	-	ADIPONITRILE
CHLOROTRIFLUOROETHYLENE	BENZOYL PEROXIDE	-	-
DIKETENE	3,4-DICHLOROANILINE	-	-
DICHLOROACETALDEHYDE	CUMENE HYDROPEROXIDE	-	-
2,3-DIHYDROFURAN	2-METHYLPROPANAL	-	-
DICUMYL PEROXIDE	2-(2-METHOXYETHOXY)ETHANOL	-	-
METHANE	1,2,3-TRICHLOROPROPANE	-	-
METHYL alpha-HYDROXYISOBUTYRATE	ACETIC ANHYDRIDE	-	-
DI-t-BUTYL PEROXIDE	DIGLYCOLIC ACID	-	-
CYCLOPENTANONE	o-NITROTOLUENE	-	-
FURAN	-	-	-
N-METHYLPYRROLE	-	-	-
tert-BUTYLAMINE	-	-	-
1,3-DIFLUOROBENZENE	-	-	-
METHACRYLONITRILE	-	-	-
2-NORBORNENE	-	-	-
FLUOROBENZENE	-	-	-
ETHYLENEIMINE	-	-	-
SALICYLALDEHYDE	-	-	-
2-METHYL-2-AMINOBUTANE	-	-	-
1,4-DIFLUOROBENZENE	-	-	-
METHYL SALICYLATE	-	-	-
PIPERIDINE	-	-	-
TETRYL	-	-	-
METHOXYDIHYDROPY	-	-	-

The changes of ADs were studied when the outliers were excluded. The changed convex hulls painted blue are shown in **Figure S11-S14**, the red wireframes depict the original ADs of model (Ⅰ). After the outliers identified by ECDF were excluded from the training and validation sets, the convex hulls representing ADs of the model (Ⅱ) were narrowed. Hereinto, more obvious changes happened on the ADs of LFL and FPT. It could be suggested that the less samples may decrease the prediction ability of the MDNN model.

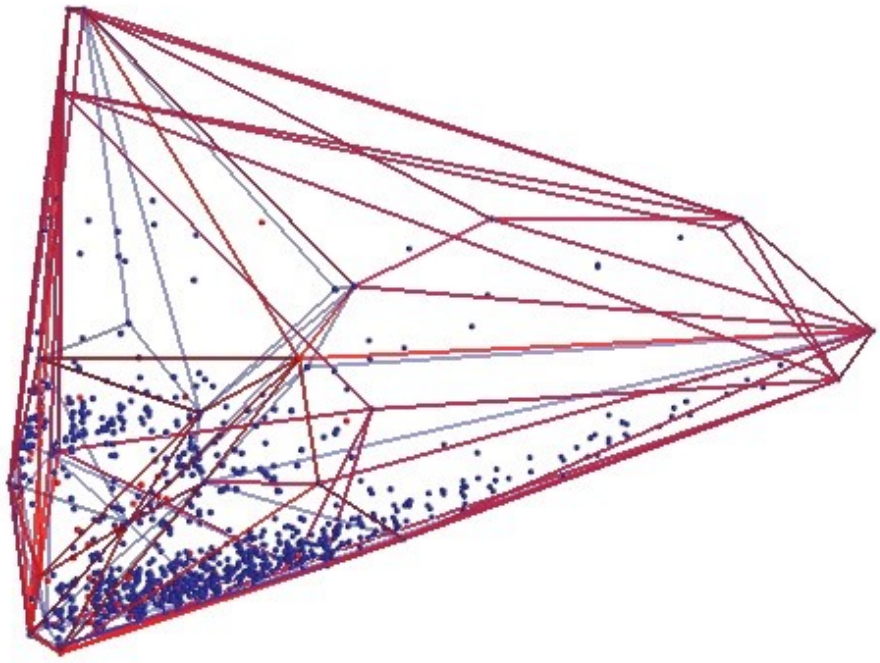

Figure S11. The convex hulls representing the changed and original ADs for FPT

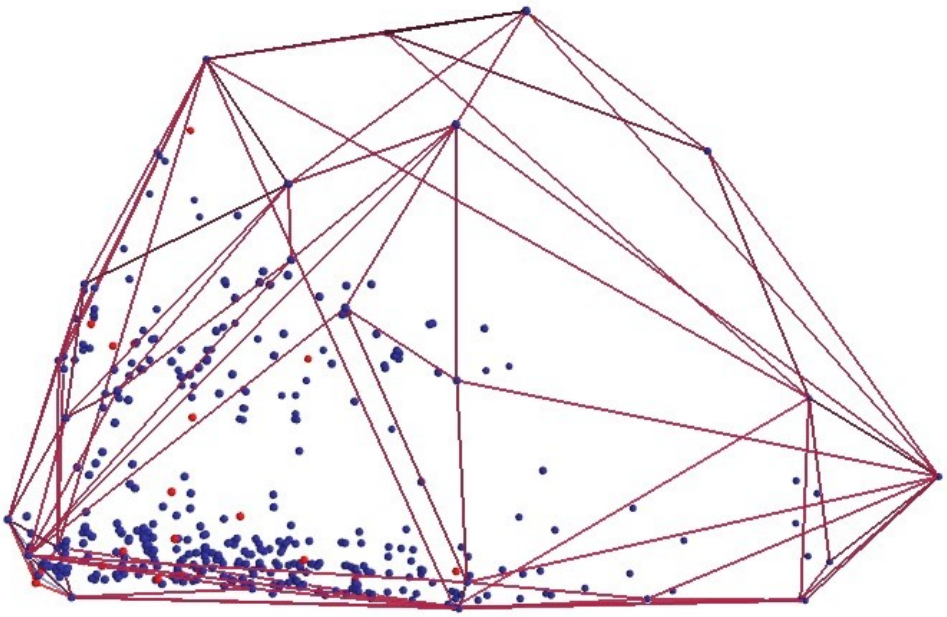

Figure S12. The convex hulls representing the changed and original ADs for AIT

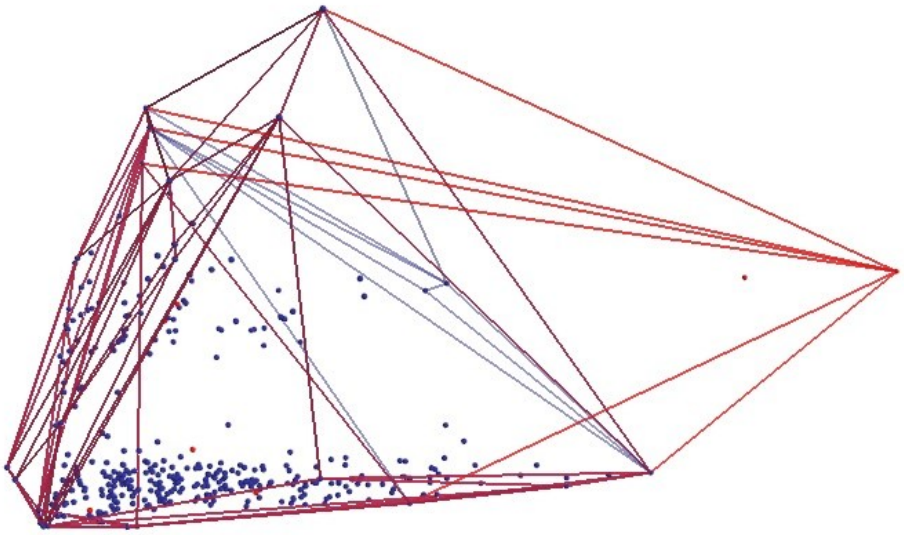

Figure S13. The convex hulls representing the changed and original ADs for LFL

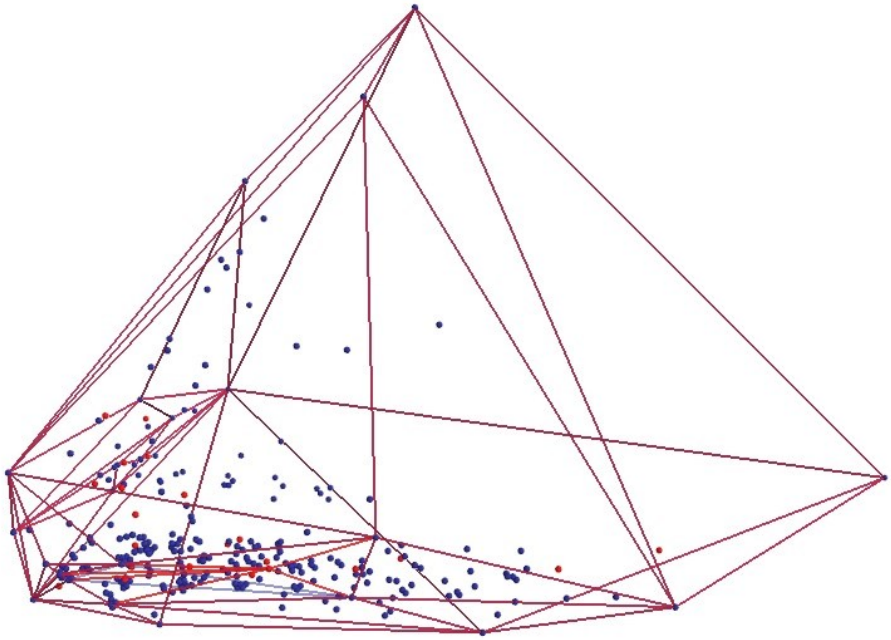

Figure S14. The convex hulls representing the changed and original ADs for UFL

We also tried to exclude the outliers from the training and validation sets of all properties and then re-trained the MDNN. As shown in **Table S11,** the obtained model performed better on the training and validation sets but only improved on test set of UFL. Even the reductions of r and R^2 could be observed on the test sets of AIT and LFL. After the samples outside the ADs, the values of r and R^2 increased for LFL significantly.

Table S11. The performance statistics of the re-trained MDNN model without all outliers

		FPT (K)	AIT (K)	LFL (%)	UFL (%)
MAE	Training set	10.15/8.995[a]	34.02/30.21[a]	0.2398/0.3207[a]	2.366/1.759[a]
	Validation set	11.49/9.695[a]	40.65/32.74[a]	0.2772/0.3066[a]	3.026/2.430[a]
	Test set	13.56/12.59[b]	56.12/54.34[b]	0.3881/0.2617[b]	2.439/2.505[b]
MPE	Training set	3.120/2.729[a]	5.367/4.519[a]	16.27/17.77[a]	18.77/13.46[a]
(%)	Validation set	3.401/2.937[a]	6.443/5.490[a]	19.87/12.10[a]	18.99/17.32[a]
	Test set	4.001/3.762[b]	9.234/8.889[b]	19.26/16.77[b]	24.15/24.45[b]
r	Training set	0.9808/0.9871[a]	0.9421/0.9204[a]	0.9754/0.9880[a]	0.9623/0.9633[a]
	Validation set	0.9643/0.9836[a]	0.9135/0.9389[a]	0.9809/0.9812[a]	0.7832/0.8608[a]
	Test set	0.9554/0.9595[b]	0.8020/0.8015[b]	0.8876/0.9736[b]	0.8752/0.8490[b]
R^2	Training set	0.9541/0.9775[a]	0.8845/0.9067[a]	0.9877/0.9586[a]	0.9238/0.9693[a]
	Validation set	0.9214/0.9616[a]	0.8173/0.8778[a]	0.9619/0.9625[a]	0.5703/0.6254[a]
	Test set	0.9022/0.9105[b]	0.5985/0.5920[b]	0.7435/0.9417[b]	0.6868/0.6107[b]

NOTE: [a] This value was obtained on the data set without the outliers identified by ECDF.

[b] This value was obtained on the data points only involved in the AD determined by training and validation sets.

S8. Distributions of compounds in various families

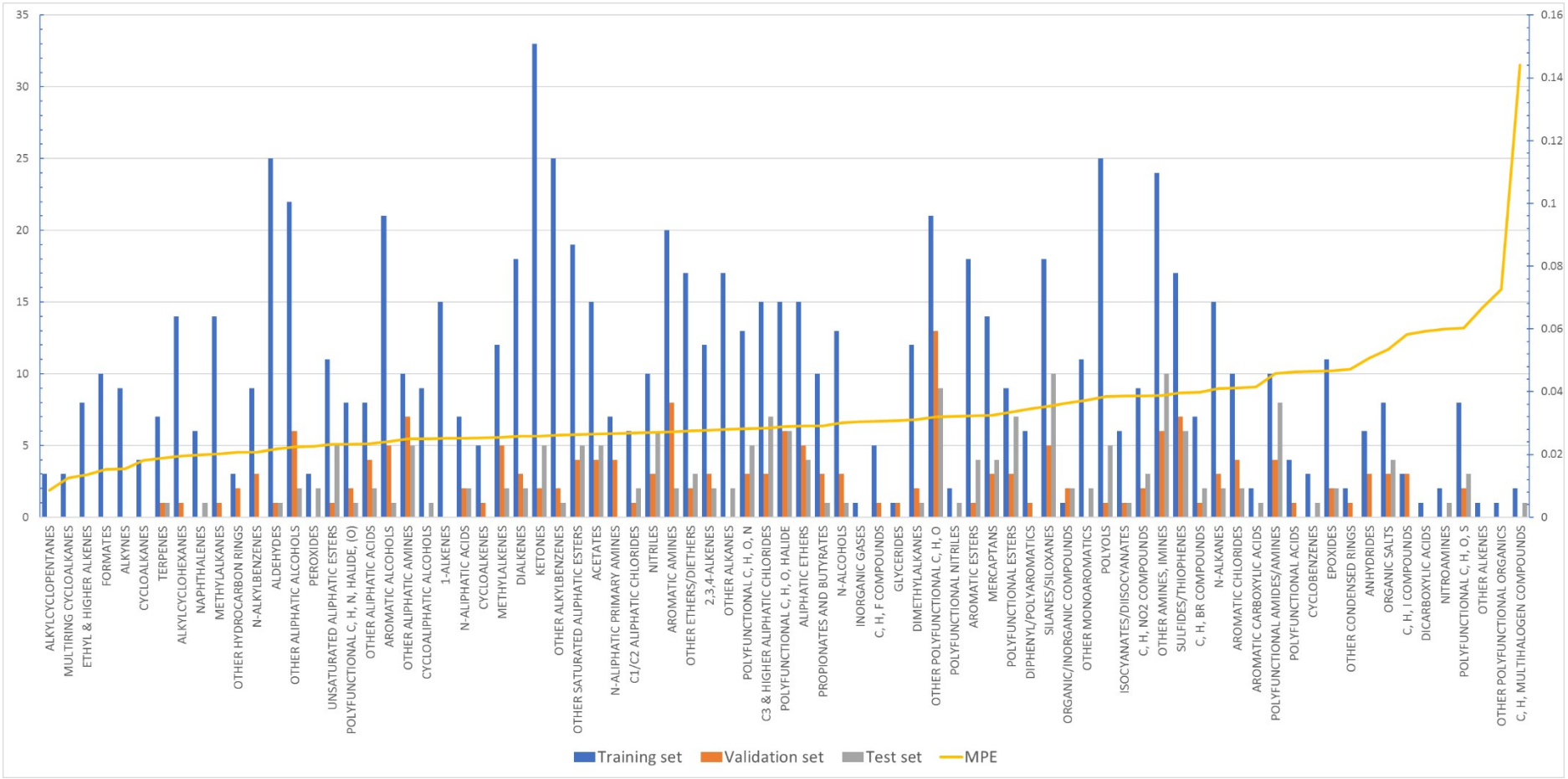

Figure S15. The distribution plot of compounds in each family with MPE for FPT

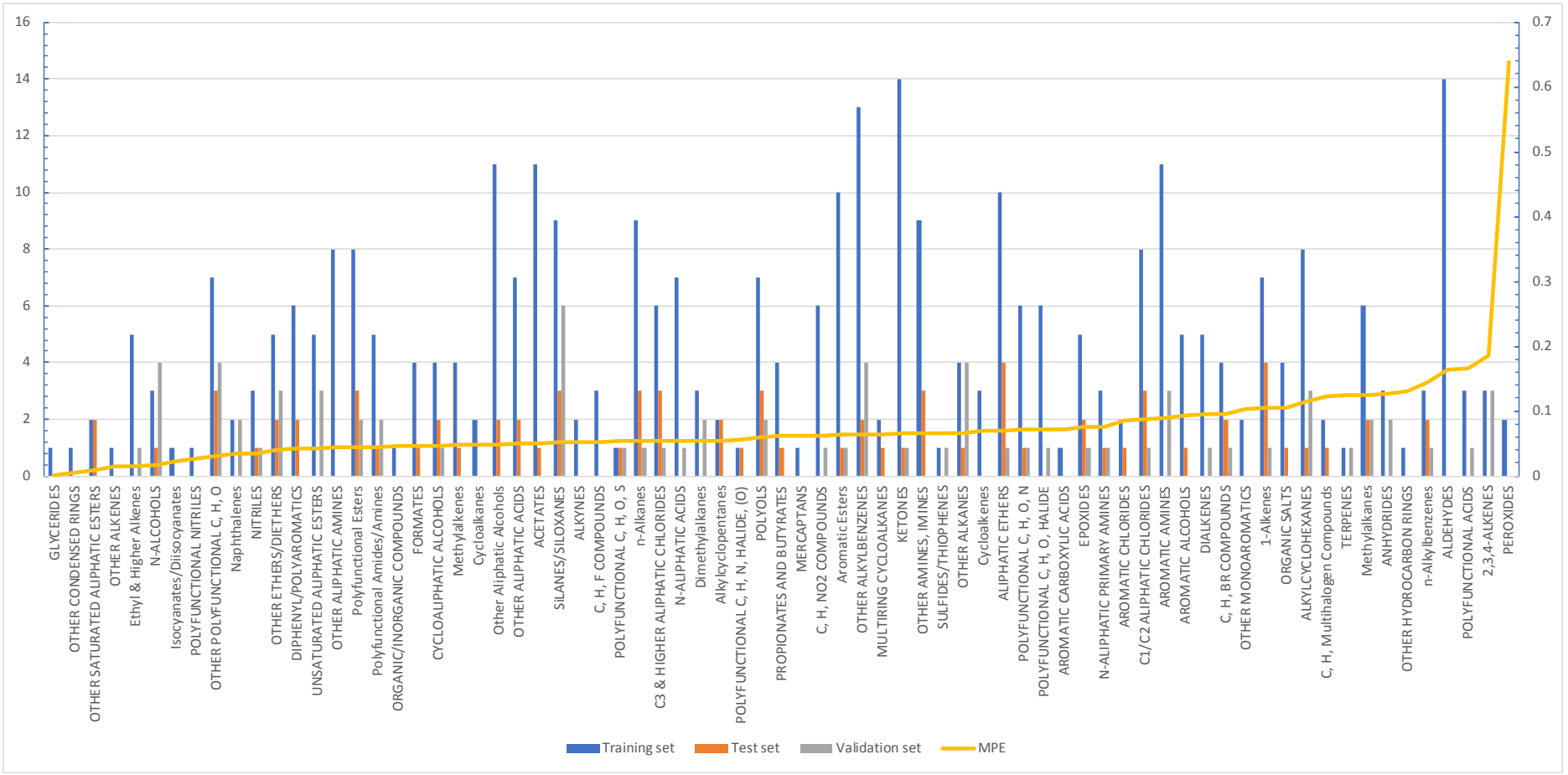

Figure S16. The distribution plot of compounds in each family with MPE for AIT

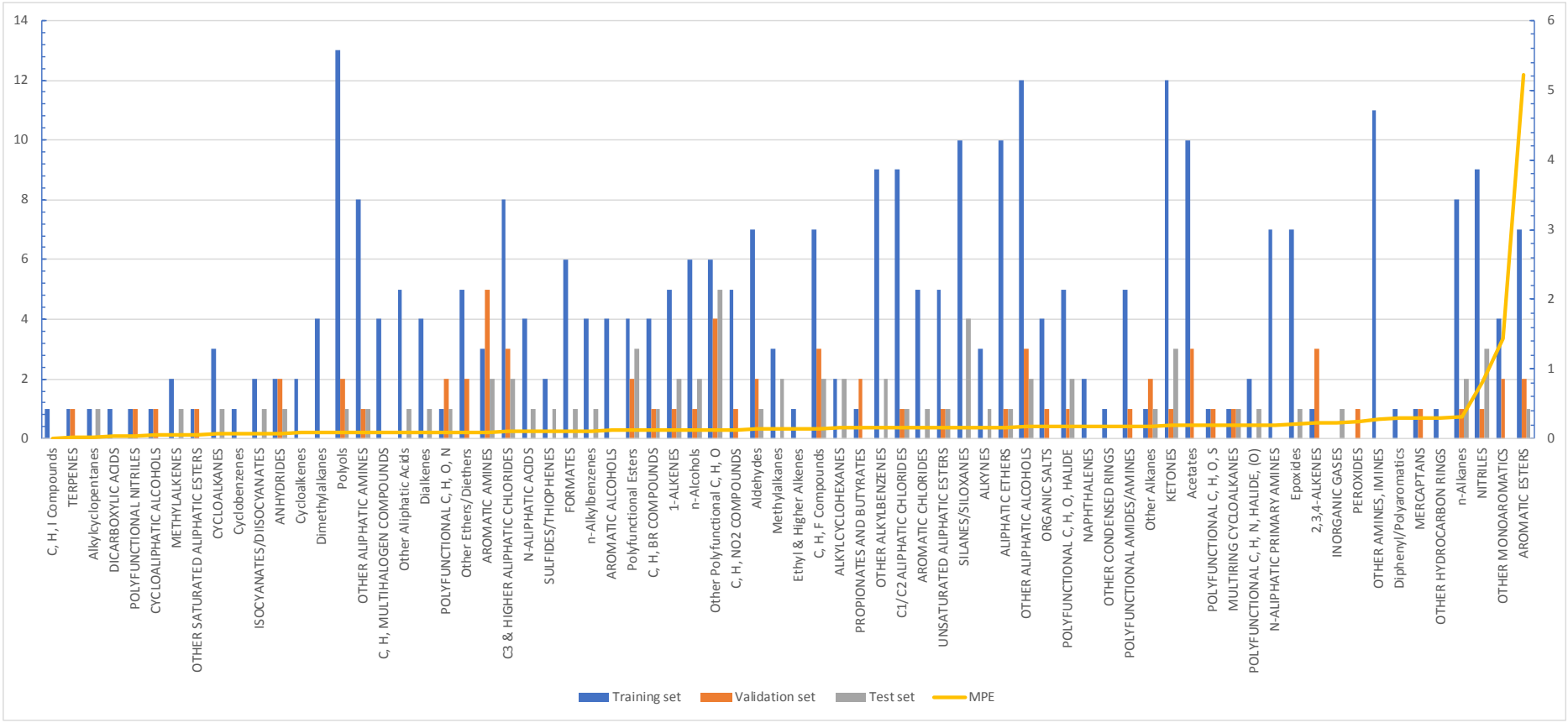

Figure S17. The distribution plot of compounds in each family with MPE for LFL

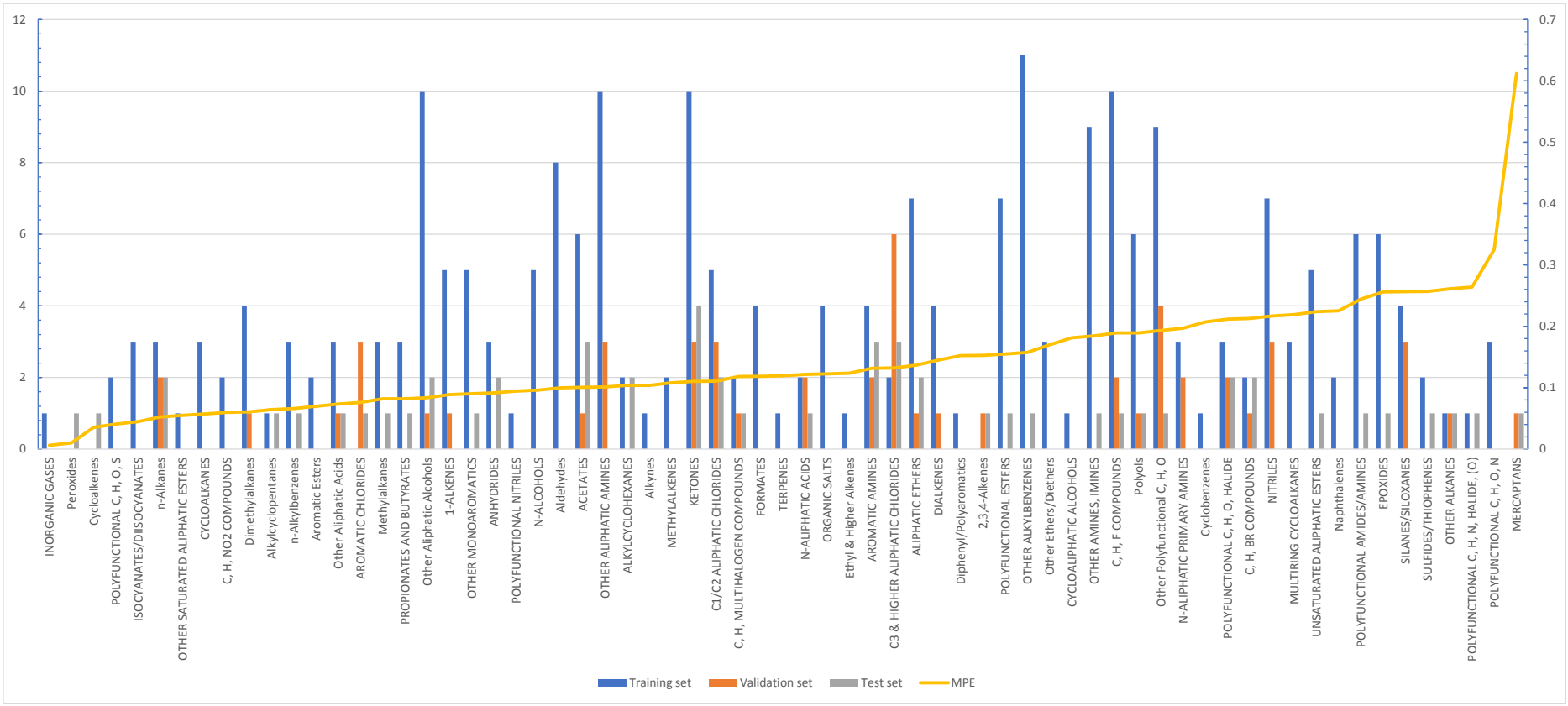

Figure S18. The distribution plot of compounds in each family with MPE for UFL

S8. Information of Software and Hardware

Software

Data analysis tool: scikit-learn[34], SciPy[37]

Deep learning framework: PyTorch[35]

Chemical information software: RDKit[36]

Data visualization: VTK[38]

Hardware

PC: Intel Core i5-8400 with 24GB RAM and Nvidia GeForce GTX 1060

S8. Acronyms

ANN	Artificial Neural Network
CNN	Convolutional Neural Network
DNN	Deep Neural Network
FNN	Feedforward neural network
GC	Group Contribution
LSTM	Long Short-Term Memory
MLR	Multiple Linear Regression
QSPR	Quantitative Structure Property Relationship
RNN	Recurrent Neural Network
SMILES	Simplified Molecular Input Line Entry System
TI	Topological Index
TIs	Topological Indices
Tree-LSTM	Tree-structured Long Short-Term Memory

S9. Notations and Symbols

c_j	cell state vector of the Tree-LSTM unit
f_{jk}	forget gate's activation vector from child nodes of the j^{th} node
h_j	output vector of the Tree-LSTM unit of the j^{th} node
$\widetilde{h}_j$	the sum of output vector from the child nodes of the j^{th} node
i_j	input gate's activation vector of the j^{th} node

N	the number of observations in the sample
o_j	output gate's activation vector of the j^{th} node
$W^{(i,u,o,f)}$, $U^{(i,u,o,f)}$, $b^{(i,u,o,f)}$	weights and biases of the formula calculating vector *i, u, o* and *f*
x_i	the observed values of the sample item *i*
x_j	input vector to the Tree-LSTM unit of the j^{th} node
x_i^{exp}	the experimental value of sample item *i*
x_i^{prep}	the prediction value of sample item *i*
$\overline{x}^{exp}$	the mean value of these observations
σ	sigmoid function
$\bullet$	multiply each element between two tensors

References

1. Faulon, J. The signature molecular descriptor. 1. using extended valence sequences in QSAR and QSPR studies. *J. Chem. Inf. Comput. Sci.* **43**, 707–720 (2003).
2. Mikolov, T., Sutskever, I., Chen, K., Corrado, G. & Dean, J. Distributed Representations of Words and Phrases and Their Compositionality. arXiv eprint. 2013;arXiv:1310.4546..
3. McCormick, C. Word2Vec tutorial-the skip-gram model. http://www. mccormickml.com. Accessed on March, 9, 2018.
4. Tai, K. S., Socher, R. & Manning, C. D. Improved semantic representations from tree-structured long short-term memory networks. *Comput Sci.* 5(1),36 (2015).
5. Lebret, R. & Collobert, R. Word Emdeddings through Hellinger PCA. arXiv eprint. 2013;arXiv:1312.5542.
6. Kim, S. *et al.* PubChem 2019 update: improved access to chemical data. *Nucleic Acids Res.* **47**, D1102–D1109 (2019).
7. Frutiger, J., Marcarie, C., Abildskov, J. & Sin, G. Group-contribution based property estimation and uncertainty analysis for flammability-related properties. *J. Hazard. Mater.* **318**, 783–793 (2016).
8. Hukkerikar, A. S. *et al.* Group-contribution+ (GC+) based estimation of properties of pure components: Improved property estimation and uncertainty analysis. *Fluid Phase Equilib.* **321**, 25–43 (2012).

9. Suzuki, T. A Method for estimating flash points of organic compounds from molecular structures. *Jounal Chem. Eng. Japan* **24**, 258–261 (1991).
10. Khajeh, A. & Modarress, H. QSPR prediction of flash point of esters by means of GFA and ANFIS. *J. Hazard. Mater.* **179**, 715–720 (2010).
11. Gharagheizi, F., Alamdari, R. F. & Angaji, M. T. A new neural network-group contribution method for estimation of flash point temperature of pure components. *Energy and Fuels* **22**, 1628–1635 (2008).
12. Pan, Y., Jiang, J., Wang, R., Cao, H. & Zhao, J. Quantitative structure-property relationship studies for predicting flash points of organic compounds using support vector machines. *QSAR Comb. Sci.* **27**, 1013–1019 (2008).
13. Albahri, T. A. MNLR and ANN structural group contribution methods for predicting the flash point temperature of pure compounds in the transportation fuels range. *Process Saf. Environ. Prot.* **93**, 182–191 (2015).
14. Patel, S. J., Ng, D. & Mannan, M. S. QSPR flash point prediction of solvents using topological indices for application in computer aided molecular design. *Ind. Eng. Chem. Res* **48**, 7378–7387 (2009).
15. Alibakhshi, A., Mirshahvalad, H. & Alibakhshi, S. Prediction of flash points of pure organic compounds: Evaluation of the DIPPR database. *Process Saf. Environ. Prot.* **105**, 127–133 (2017).
16. Suzuki, T. Quantitative structure-property relationships for auto-ignition temperatures of organic compounds. *Fire Mater.* **18**, 81–88 (1994).
17. Gharagheizi, F. An accurate model for prediction of autoignition temperature of pure compounds. *J. Hazard. Mater.* **189**, 211–221 (2011).
18. Albahri, T. A. & George, R. S. Artificial neural network investigation of the structural group contribution method for predicting pure components auto ignition temperature. *Ind. Eng. Chem. Res.* **42**, 5708–5714 (2003).
19. Pan, Y., Jiang, J., Wang, R., Cao, H. & Cui, Y. Predicting the auto-ignition temperatures of organic compounds from molecular structure using support vector machine. *J. Hazard. Mater.* **164**, 1242–1249 (2009).
20. Pan, Y., Jiang, J., Wang, R., Cao, H. & Zhao, J. Prediction of auto-ignition temperatures

of hydrocarbons by neural network based on atom-type electrotopological-state indices. *J. Hazard. Mater.* **157**, 510–517 (2008).

21. Design Institute for Physical Properties. Design Institute for Physical Properties, Sponsored by AIChE. (2005; 2008; 2009; 2010; 2011; 2012; 2015; 2016; 2017; 2018; 2019). DIPPR Project 801 - Full Version.Knovel.

22. Rowley, J. R., Rowley, R. L. & Wilding, W. V. Estimation of the lower flammability limit of organic compounds as a function of temperature. *J. Hazard. Mater.* **186**, 551–557 (2011).

23. Seaton, W. H. Group contribution method for predicting the lower and the upper flammable limits of vapors in air. *J. Hazard. Mater.* **27**, 169–185 (1991).

24. Mendiburu, A. Z., de Carvalho Jr, J. A. & Coronado, C. R. Estimation of lower flammability limits of CH compounds in air at atmospheric pressure, evaluation of temperature dependence and diluent effect. *J. Hazard. Mater.* **285**, 409–418 (2015).

25. Albahri, T. A. Flammability characteristics of pure hydrocarbons. *Chem. Eng. Sci.* **58**, 3629–3641 (2003).

26. Lazzús, J. A. Neural network/particle swarm method to predict flammability limits in air of organic compounds. *Thermochim. Acta* **512**, 150–156 (2011).

27. Gharagheizi, F. Quantitative structure-property relationship for prediction of the lower flammability limit of pure compounds. *Energy & Fuels* **22**, 3037–3039 (2008).

28. Gharagheizi, F. A new group contribution-based model for estimation of lower flammability limit of pure compounds. *J. Hazard. Mater.* **170**, 595–604 (2009).

29. Pan, Y., Jiang, J., Wang, R., Cao, H. & Cui, Y. A novel QSPR model for prediction of lower flammability limits of organic compounds based on support vector machine. *J. Hazard. Mater.* **168**, 962–969 (2009).

30. Gharagheizi, F. Prediction of upper flammability limit percent of pure compounds from their molecular structures. *J. Hazard. Mater.* **167**, 507–510 (2009).

31. Pan, Y., Jiang, J., Wang, R., Cao, H. & Cui, Y. Prediction of the upper flammability limits of organic compounds from molecular structures. *Ind. Eng. Chem. Res.* **48**, 5064–5069 (2009).

32. High, M. S. & Danner, R. P. Prediction of upper flammability limit by a group

contribution method. *Ind. Eng. Chem. Res.* **26**, 1395–1399 (1987).

33. Mendiburu, A. Z., de Carvalho, J. A. & Coronado, C. R. Determination of upper flammability limits of CHO compounds in air at reference temperature and atmospheric pressure. *Fuel* **188**, 212–222 (2017).

34. Pedregosa, F. *et al.* Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2013).

35. Paszke, A., Gross, S., Chintala, S. & Chanan, G. Pytorch: Tensors and dynamic neural networks in python with strong gpu acceleration. https://pytorch.org. Accessed on January, 23, 2019.

36. Landrum, G. RDKit: Open-source cheminformatics software. http://www.rdkit.org/, https://github.com/rdkit/rdkit. Accessed on March, 9, 2019.

37. The SciPy community. SciPy, a Python-based ecosystem of open-source software for mathematics, science, and engineering. https://docs.scipy.org/doc/scipy/reference/generated/scipy.spatial.ConvexHull.html. Accessed on May, 6, 2020.

38. Kitware. The Visualization Toolkit (VTK), https://vtk.org/. Accessed on March, 20, 2020.