Gilmer et al. (2017)

Neural Message Passing for Quantum Chemistry

Andreas Holm (HDF957) & Tim Mondorf (CJK681) 30 September 2025

UNIVERSITY OF COPENHAGEN

References (A)

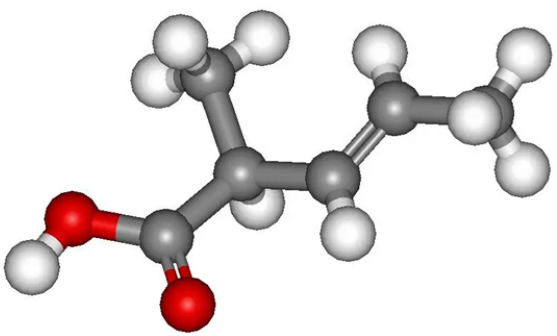
- Main article: *Justin Gilmer, Samuel S. Schoenholz, Patrick F. Riley, Oriol Vinyals, George E. Dahl Neural Message Passing for Quantum Chemistry, International Conference on Machine Learning (2017)*
- Simpler model based on Gilmer: *Bishop and Bishop Deep Learning Foundations and Concepts (2024), chapter 13 Graph Neural Networks*
- QM9 Dataset used by Gilmer: *Raghunathan Ramakrishnan, Pavlo Dral, Matthias Rupp and Anatole Lilienfeld: Quantum chemistry structures and properties of 134 kilo molecules, Scientific Data (2014)*

The simple model of Bishop (2024) (T)

- Supervised learning, observations are molecules, classified as water soluble/insoluble
- Applying ML to a deterministic rather than stochastic process
- Nodes, attributes X, edges, adjacency matrix A

$$X = \begin{bmatrix} x_{11}, \ldots, x_{1D} \\ \ldots \\ x_{N1}, \ldots, x_{ND} \end{bmatrix}$$

$$A = \begin{bmatrix} 0, \ldots 1 \\ \ldots \\ 1, \ldots, 0 \end{bmatrix}$$

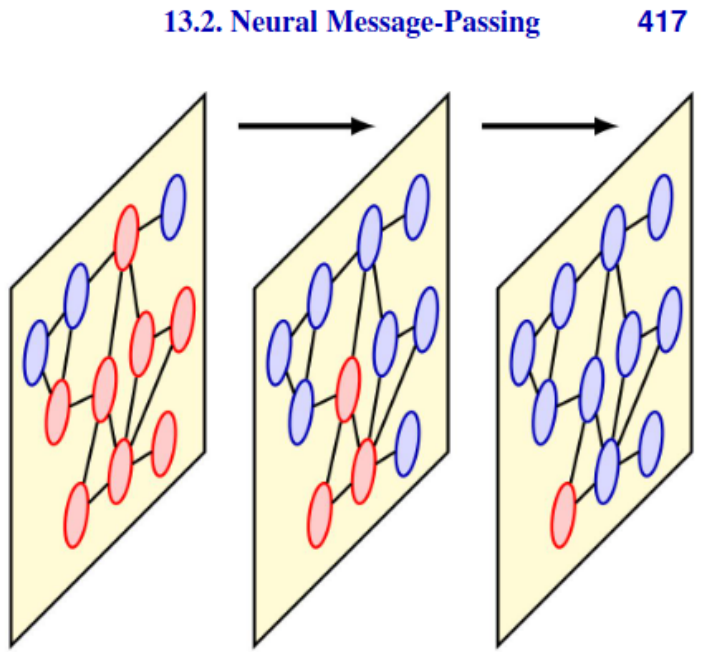


From Chantiers de Science https://jcmarot.com/2019/02/03/molecules-1-isomeries/

- i.i.d. is out and permutation equivariance is in

But ... diffferent molecules have different dimensionality? (A)

- Convolution is your friend
- Node (atom) x_v has a hidden value h_v that is updated by each layer
- $h_v^{t+1} = Update(h_v^t, Aggregate(h_m^t : m \in N(x_v))$
- The larger the molecule, the deeper the network



Bishop, Figure 13.4

Gilmer (2017) (T)

- Predict not one but 13 different variables related to molecules
- Two types of known variable values

  1. What chemists already know (does not cover all 130.000 molecules)
  2. What can be approximated using quantum computing (expensive)

- ML is almost as good as (1) and much less expensive than (2)

No.	Property	Unit	Description
1	tag	—	'gdb9' string to facilitate extraction
2	i	—	Consecutive, 1-based integer identifier
3	A	GHz	Rotational constant
4	B	GHz	Rotational constant
5	C	GHz	Rotational constant
6	μ	D	Dipole moment
7	α	a_0^3	Isotropic polarizability
8	ϵ_{HOMO}	Ha	Energy of HOMO
9	ϵ_{LUMO}	Ha	Energy of LUMO
10	ϵ_{gap}	Ha	Gap ($\epsilon_{LUMO} - \epsilon_{HOMO}$)
11	$\langle R^2 \rangle$	a_0^2	Electronic spatial extent
12	zpve	Ha	Zero point vibrational energy
13	U_0	Ha	Internal energy at 0 K
14	U	Ha	Internal energy at 298.15 K
15	H	Ha	Enthalpy at 298.15 K
16	G	Ha	Free energy at 298.15 K
17	C_v	$\frac{cal}{molK}$	Heat capacity at 298.15 K

Table 3. Calculated properties. Properties are stored in the order given by the first column.

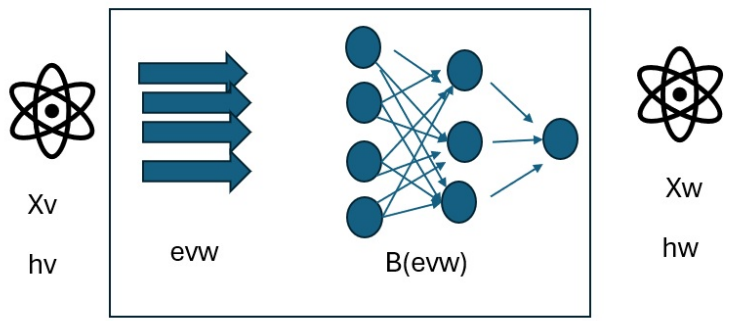
Ramakrishnan (2014)

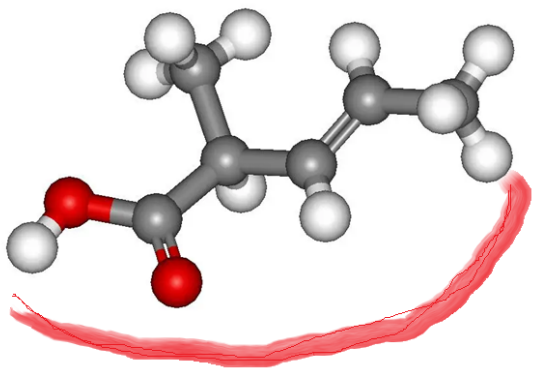
Predict one variable (A)

- Edges become multi-dimensional real-valued vectors
- Message from atom x_v to x_w

$$M(h_v, h_w, e, vw) = B(e_{vw})h_w$$

- Notation suppresses number of layer
- $O(n^2d^2)$ - but large molecules are rarely cliques
- If the largest molecule in the training data is big, build a deep network and vice versa
- Not robust if test molecules are much larger than those in the training data (no i.i.d.)
- Next: 'attention mechanism' across all atoms





Gilmer Table 2: Results (T)

Table 2. Comparison of Previous Approaches (left) with MPNN baselines (middle) and our methods (right)

	Previous non-graph models					Previous graph models			Gilmer	
									Graph model	Ensemble
Variables: Target	BAML	BOB	CM	ECFP4	HDAD	GC	GG-NN	DTNN	enn-s2s	enn-s2s-ens5
mu	4.34	4.23	4.49	4.82	3.34	0.70	1.22	-	**0.30**	0.20
alpha	3.01	2.98	4.33	34.54	1.75	2.27	1.55	-	**0.92**	0.68
HOMO	2.20	2.20	3.09	2.89	1.54	1.18	1.17	-	**0.99**	0.74
LUMO	2.76	2.74	4.26	3.10	1.96	1.10	1.08	-	**0.87**	0.65
gap	3.28	3.41	5.32	3.86	2.49	1.78	1.70	-	**1.60**	1.23
R2	3.25	0.80	2.83	90.68	1.35	4.73	3.99	-	**0.15**	0.14
ZPVE	3.31	3.40	4.80	241.58	1.91	9.75	2.52	-	**1.27**	1.10
U0	1.21	1.43	2.98	85.01	0.58	3.02	0.83	-	**0.45**	0.33
U	1.22	1.44	2.99	85.59	0.59	3.16	0.86	-	**0.45**	0.34
H	1.22	1.44	2.99	86.21	0.59	3.19	0.81	-	**0.39**	0.30
G	1.20	1.42	2.97	78.36	0.59	2.95	0.78	.84[2]	**0.44**	0.34
Cv	1.64	1.83	2.36	30.29	0.88	1.45	1.19	-	**0.80**	0.62
Omega	0.27	0.35	1.32	1.47	0.34	0.32	0.53	-	**0.19**	0.15
Average	2.17	2.08	3.37	53.97	1.35	2.59	1.36	-	**0.68**	0.52

- Unit is proportion of the error that chemists make using traditional methods
- ChatGPT was used to mark all errors green if <1