

Predicting Molecular Properties

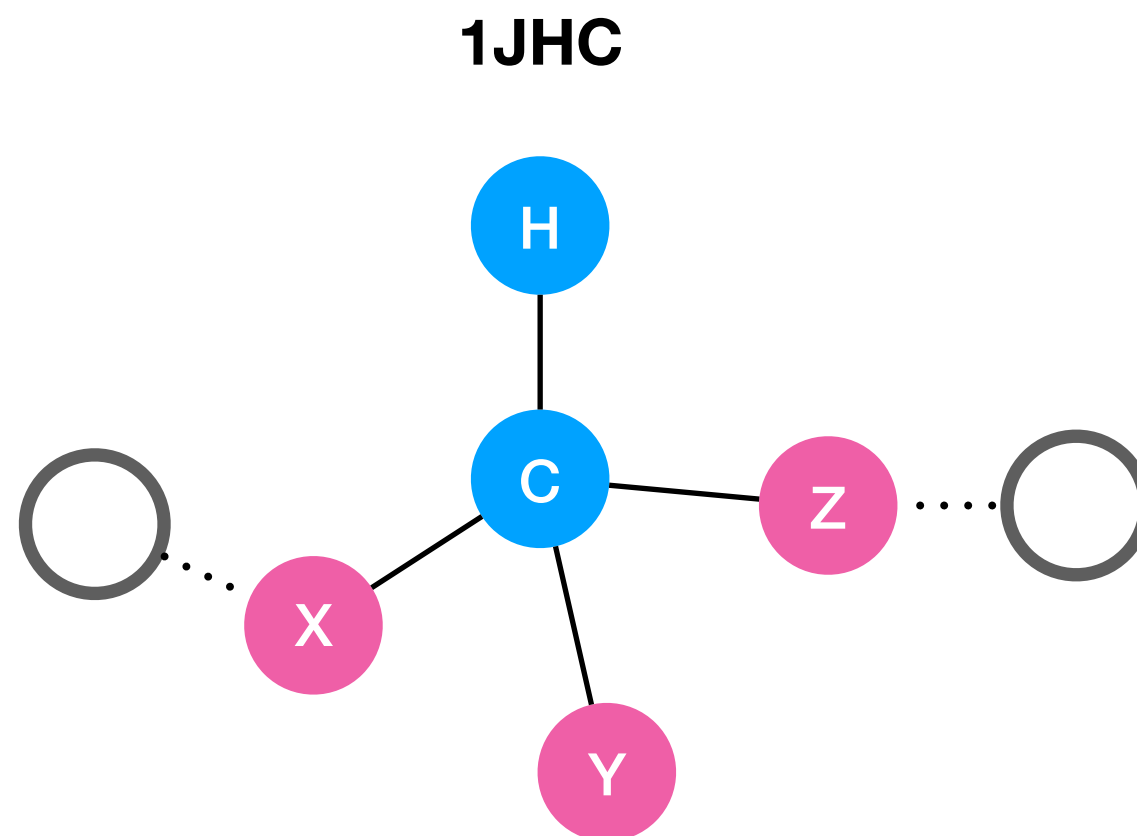
10 th solution, “Kha Zidmie Josh Kyle Akira” Team
Akira Sosa

Challenging points

- Graph (Brand new competition in Kaggle)
- Some extra data such as magnetic shielding tensors
- Different kinds of type

Graph

Why lgb can't beat GCN?



**Even if you make features for nodes X, Y and Z,
HOW do you assign them to columns?**

**They are not distinct between different
molecules.**

Aggregate? -> lose detail

Why lgb can't beat GCN?

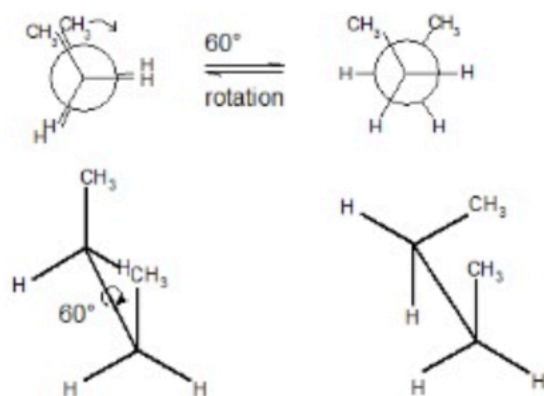
- The ability of NN for multitask learning. We have some extra data.

At last, we have given up on lgb at the very beginning.

Which GCN to use?

- 既存手法の限界

- ノード間の相互作用をグラフ上での距離のみに基づいてモデリングしている
 - グラフ構造は同じでも異なる配置が存在し、それにより予測対象の傾向が変わる場合を想定していない
 - グラフ上の距離は遠いが実空間では近いような場合に、相互作用を適切にモデリングできない可能性がある



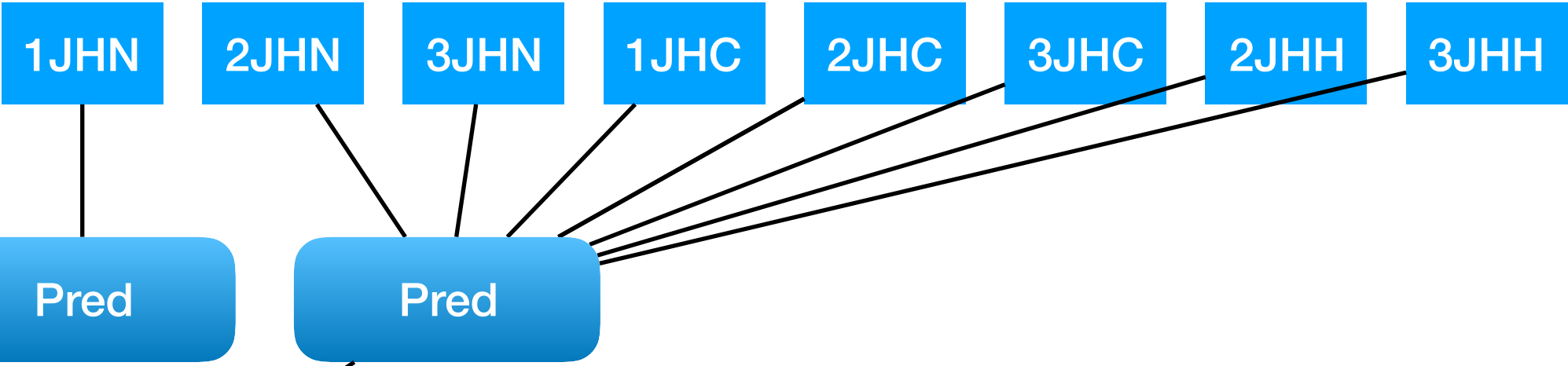
図引用: https://en.wikipedia.org/wiki/Conformational_isomerism

<https://www.slideshare.net/KazukiFujikawa/schnet-a-continuousfilter-convolutional-neural-network-for-modeling-quantum-interactions>

The answer is clear. We have already seen some geometric features such as dihedral are really important. Only **SchNet** can fully play with them.

Pipeline

Model1 for each types



Model2 for each types



-2.6

-2.8

-2.9

1JHN Pred

Model1 detail

- SchNet
 - Concatenate node features and predict atom pairwise
- Input
 - Mulliken charges
 - Dihedral Graph (next page)
- Output
 - Scalar coupling constant and 4 contributions
 - 3 atomwise values originated from magnetic shielding tensor
- Multitask learning
- Fine tune from QM9 energy (used only for 1JHN)
- Snapshot ensemble using CLR

Dihedral Graph

Dist Mat

0, 4, 1, 3, 3
4, 0, 3, 1, 1
1, 3, 0, 1, 1
3, 1, 1, 0, 1
3, 1, 1, 1, 0



Adjacency (dist=3)

0, 0, 0, 1, 1
0, 0, 1, 0, 0
0, 1, 0, 0, 0
1, 0, 0, 0, 0
1, 0, 0, 0, 0

Edge feats

- * $\cos(\text{dihedral})$
- * $\cos(2 * \text{dihedral})$

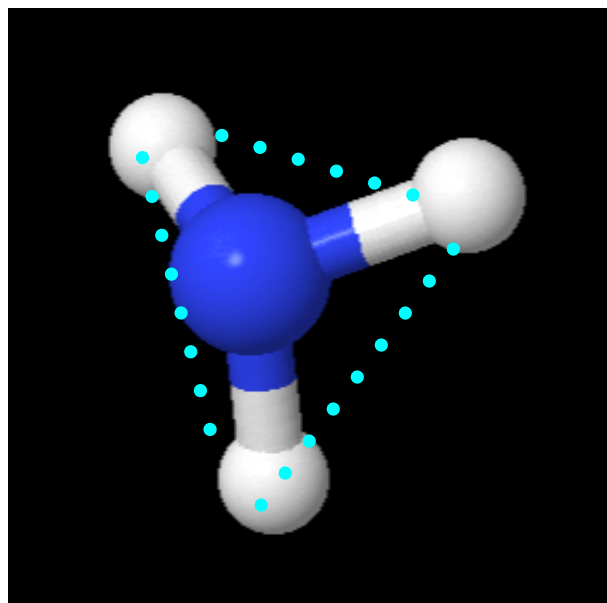
As SchNet does not have explicit edge features, I converted this edge feature to node features by NNConv.

Model2

What and Why model2?

Shortcoming of Model1

1JHN



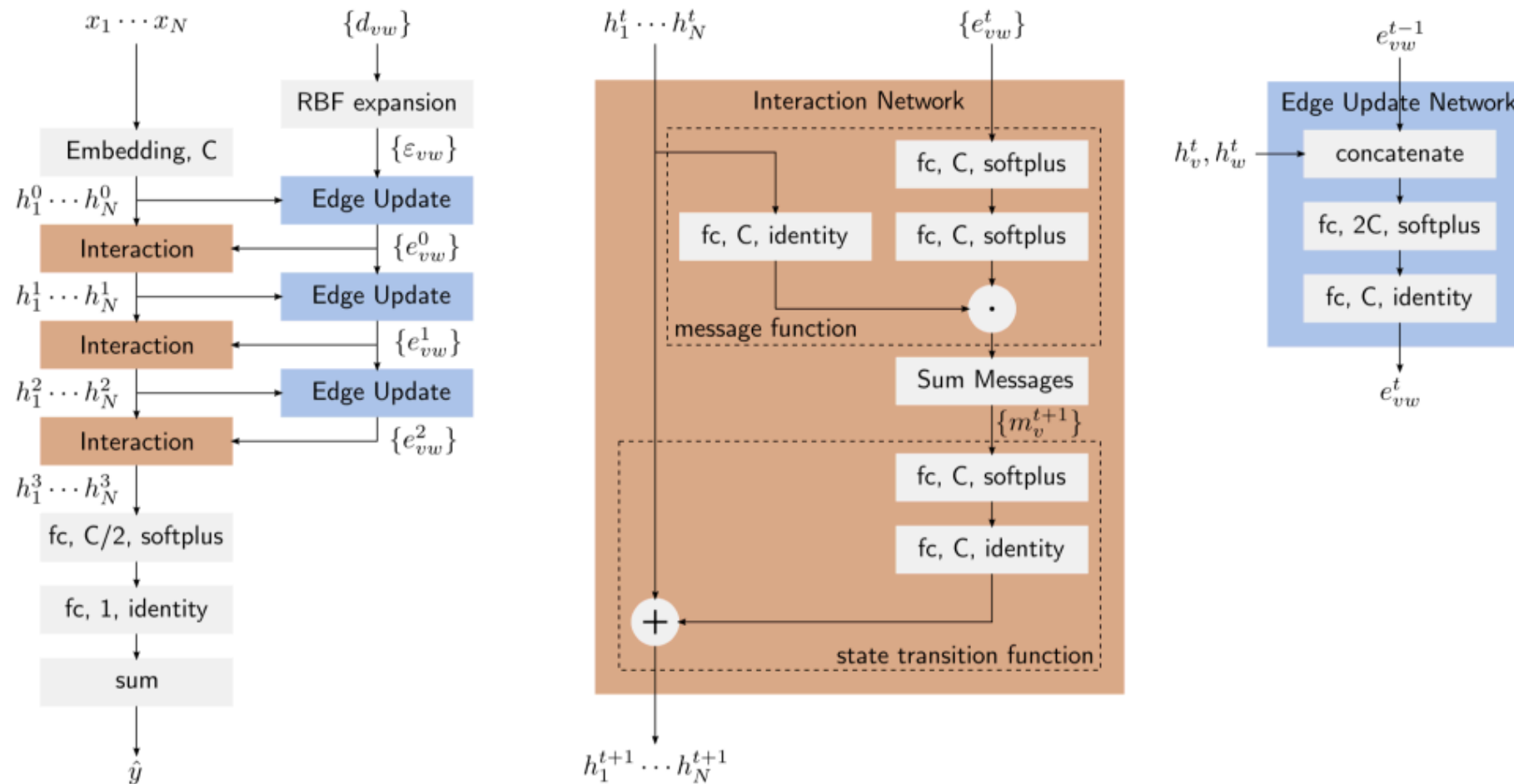
Type	atom_index0	atom_index1	Scalar coupling constant
1JHN	1	0	
1JHN	2	0	
2JHH	1	2	
2JHH	2	3	

These information are not used in training 1JHN model1.

Model2 detail

- SchNet with Edge Update (SchNet-EU)
 - Concatenate node1, node2 and edge12 features and predict atom pairwise
- Input
 - Mulliken charges
 - Edge features
 - Bond existence (dist=1 or not)
 - Angle radian
 - Angle existence (dist=2 or not)
 - Dihedral cos
 - Dihedral cos2
 - Dihedral existence (dist=3 or not)
 - **Scalar coupling constant and contributions except target type**
 - **Scalar coupling types (one hot encoded) except target type**
- Output
 - Scalar coupling constant and contributions
 - 3 atomwise values originated from magnetic shielding tensor

SchNet EU



It has **explicit edge features** so that I can feed edge features as an input more efficiently. It's crucial important especially for feeding scalar coupling constant and contributions as an input.

What did not work

- Multitask learning with ...
 - Energy
 - Dipole
- Huber, Logcosh and other losses
 - Log L1 was the best (no idea why)
- Training with noise
 - Dropout
 - DropEdge
- SchNOrb (Unifying machine learning and quantum chemistry – a deep neural network for molecular wavefunctions)
 - I used it partially to get edge feature.
 - It does not change so much from SchNet and slow.

Thank you!