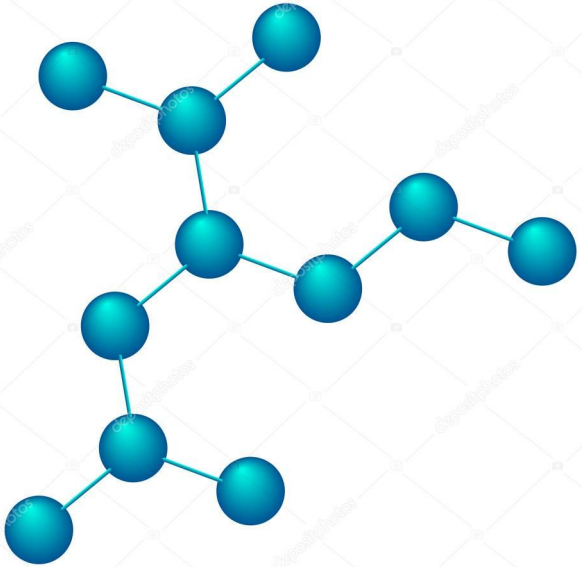
Molecular Property Prediction



Hyunwook Paul Shin

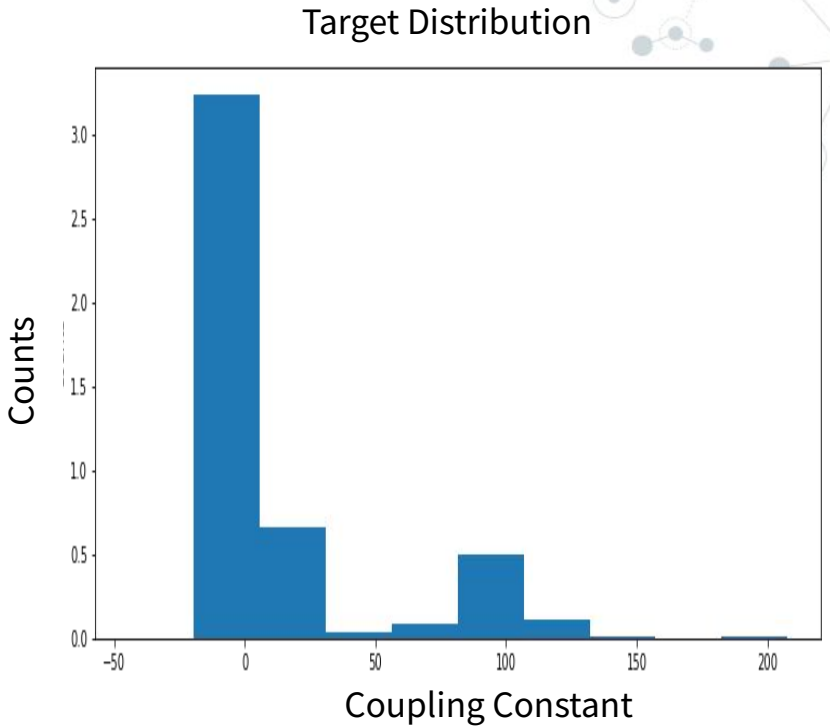
Business Problem

- Nuclear Magnetic Resonance (NMR)
  - Protein Target Drug design
- Coupling Constant
  - Limitation in Calculation

Create a model that accurately predicts coupling constant in a small amount of time.

Data

- Kaggle Competition
- 85k Unique Molecules
- 3D Structure Information
- Imbalanced Target Variable
- 8 Different Coupling Types



Results

- Neural Network
- 94.6% variance explained
- Average error = 1.799
- Over 931k predictions
  - 90 seconds

Conclusion

This model can be used to:

- Predict Coupling Constant
- Save time in interpreting NMR results

Next Steps

- Additional Features
- More Molecule Samples
- Divide into Subset

Thank You!

Any questions?

Project Github:

https://github.com/hps1795/Molecular_Properties_Prediction

Email:

paulshin17@gmail.com