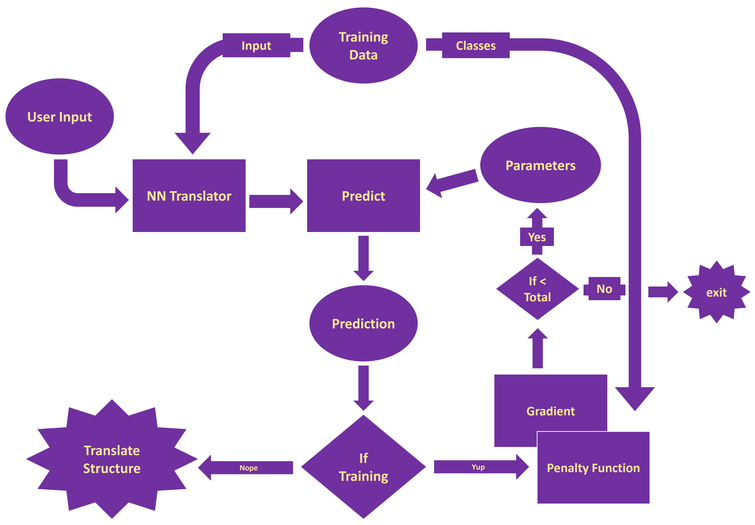
Spectral Structure Prediction In Neural Networks

Lauren Koulias, Ian Murphy, Torin Stetina, Andrew Wildman

Overview

Software package using neural networks to predict the structure of simple organic molecules based on nuclear magnetic resonance data



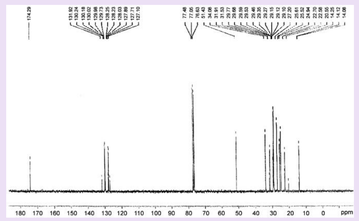
Use Cases

Automatic structural prediction from ^{13}C NMR

Train custom nets using modular framework provided by package

Visualize arbitrary chemical structures from connectivity information

Data Input

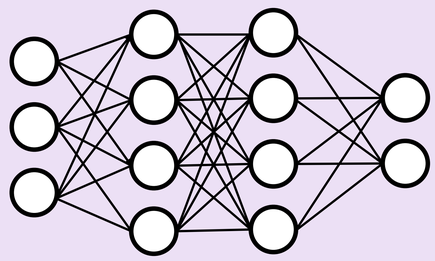


EMPIRICAL FORMULA		
SHIFT	PEAK AREA	MULTIPLICITY
11.3	1	S
21.5	1	D
22.7	2	T
.....		

Export experimental data to a space delimited .txt file separating chemical shift, intensity, and multiplicity.

Prediction

Our neural network predicts the bonding behavior between atoms in the form of a connectivity matrix

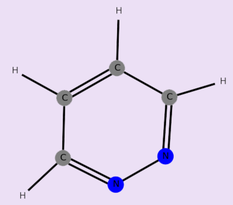


<u>Architecture</u>

- 3,350 → 1,000 ; ELU
- 1,000 → 3,000 ; Sigmoid
- 3,000 → 186,624 ; ReLU

Visualization

C	C	C	C	N	N	H	H
0	1	0	0	0	0	1	0
0	0	1	0	0	1	1	0
1	0	1	0	2	0	0	0
0	0	0	1	0	2	0	1
2	1	1	0	1	0	0	1
.....							



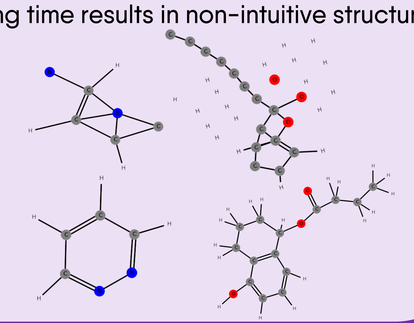
The full connectivity matrix is truncated to only include atoms which are predicted to participate in bonding with neighbors
This smaller matrix is then plotted using the NetworkX visualization package

Results

- Net outputs a predicted connectivity matrix
- Bonding thresholded to integer values
- Limited training time results in non-intuitive structures

Predicted

Real



Future Work

Challenges	Potential Solutions
Extremely sparse output	-Take upper-triangle of connectivity matrix
Large output space	-Change output space -Bonding network -HOSE
^{13}C NMR carries less than ideal structural information	-Database with ^{1}H NMR multiplicities
Functional group context	-Add more specific spectroscopy -Change output space
Sub-optimal prediction	-Increased neural net training time

References

NMR Database:
http://nmrshiftdb.nmr.uni-koeln.de
Packages:
Keras, TensorFlow, Numpy, NetworkX

DIRECT
Data Intensive Research Enabling Clean Technologies

CLEAN ENERGY INSTITUTE
UNIVERSITY of WASHINGTON

SSPINN
https://github.com/awild82/SSPINN

UNIVERSITY of WASHINGTON