Transforming the Analysis of Spectroscopic Data with Machine Learning

Julia Kaczmarek, Tom Penfold
j.a.kaczmarek2@newcastle.ac.uk
School of Natural and Environmental Sciences, Stage 3 MChem Chemistry (Hons)

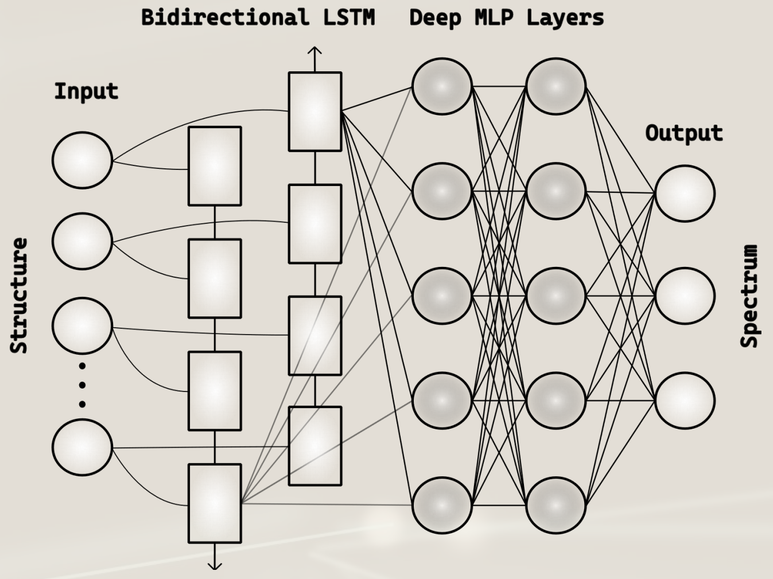
Background

~$ Time-resolved spectroscopy (measurement of time-dependent interactions between light and molecules) is the main experimental tool for probing quantum nature of matter. Complex calculations are required to understand the data - machine learning can make this less complex. Advances in this field are essential in developing novel sustainable solutions to global problems.

~$ XANESNET is a machine learning program that predicts x-ray spectra. The original neural network model does this well, but struggles to predict the effect of structural distortions, which are present in time-resolved data.

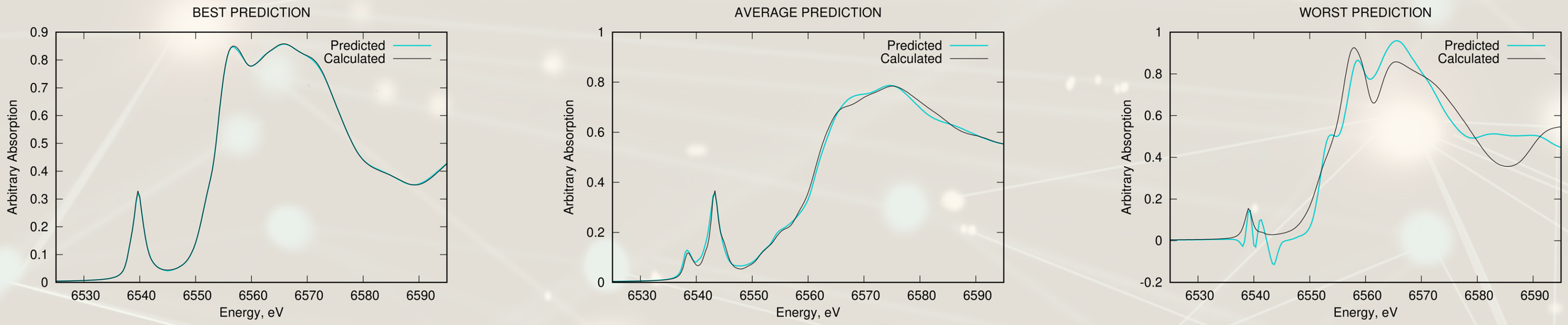
~$ We wanted to test another model, called LSTM (Long Short-Term Memory) to see whether it was capable of solving this issue.

Model



Results

~$ LSTM predicts spectra with the same or better accuracy than the original model.



~$ Creating new training data using computational chemistry methods is time-consuming and requires substantial computational power - we used the network to expand the training data:

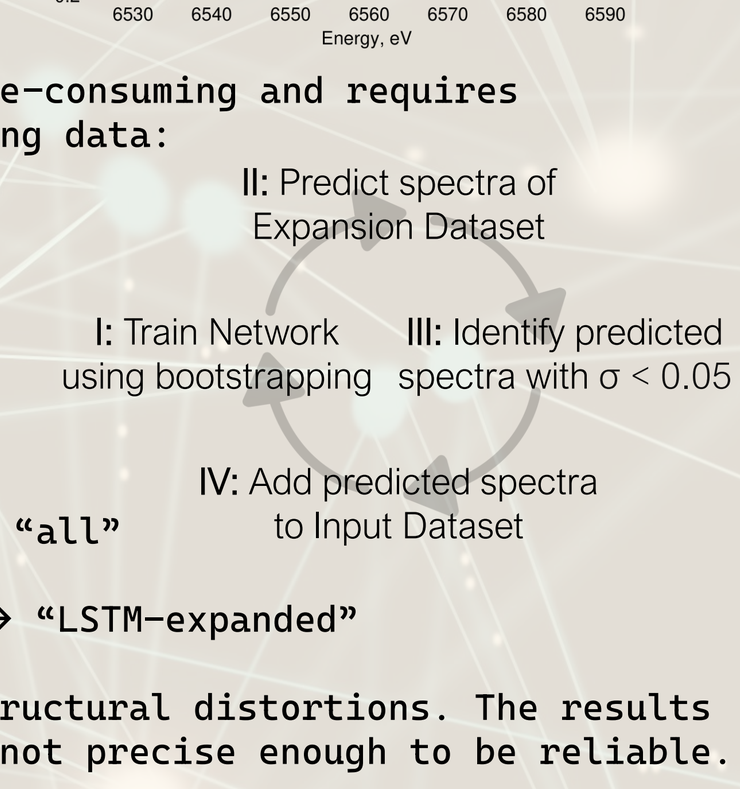
Four different training sets were tested:

I. ~1.5k Mn samples from the database → "original"

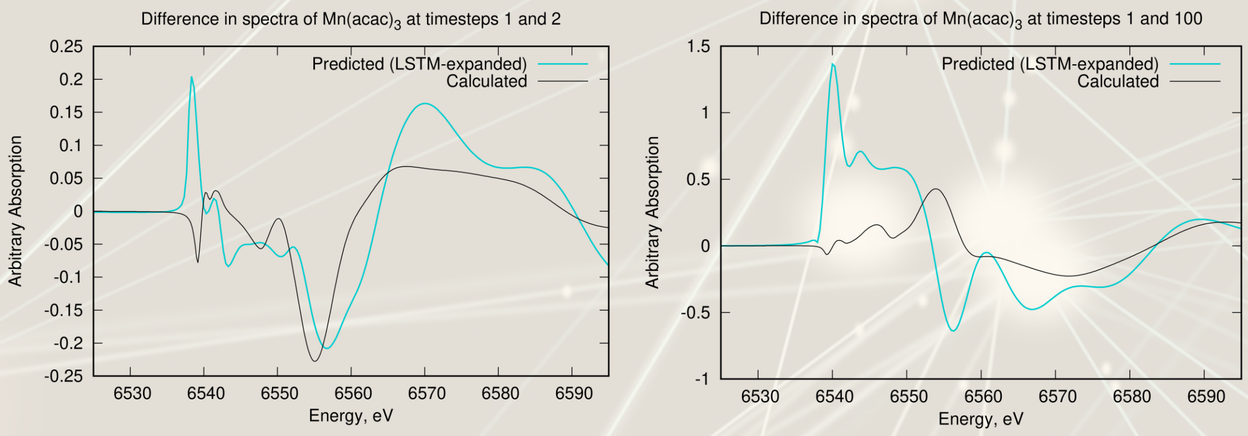
II. ~75k "original" expanded with molecular dynamics and the XANESNET MLP with standard deviation < 0.05 → "expanded"

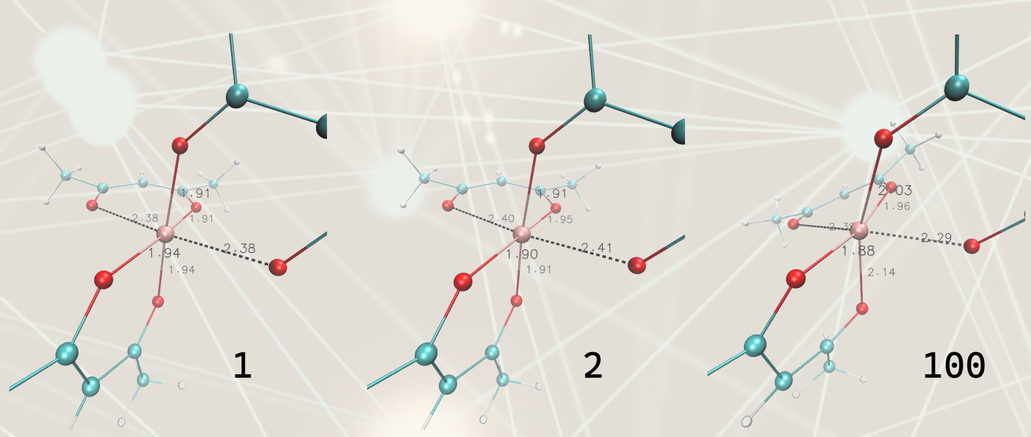
III. ~30k samples from the database for all first-row transition metals → "all"

IV. ~20k Mn samples obtained with the LSTM and standard deviation < 0.03 → "LSTM-expanded"



~$ Two manganese complexes were used as case studies for sensitivity to structural distortions. The results showed an improvement over the original network although the results were not precise enough to be reliable.





Conclusions

~$ The LSTM model has been integrated into the XANESNET code.
~$ Further work is needed to enhance the structural sensitivity.

References

1. C. D. Rankine, and T. J. Penfold, *J. Chem. Phys.*, 2022, 156, 164102
2. C. D. Rankine, M. M. M. Madkhali, and T. J. Penfold, J. Phys. Chem. A, 2020, 124, 4263-4270.

Acknowledgements

Penfold Group is the best. I would to thank Tom for being a great mentor and the University for both the funding and the opportunity.

Scan to see the XANESNET code: