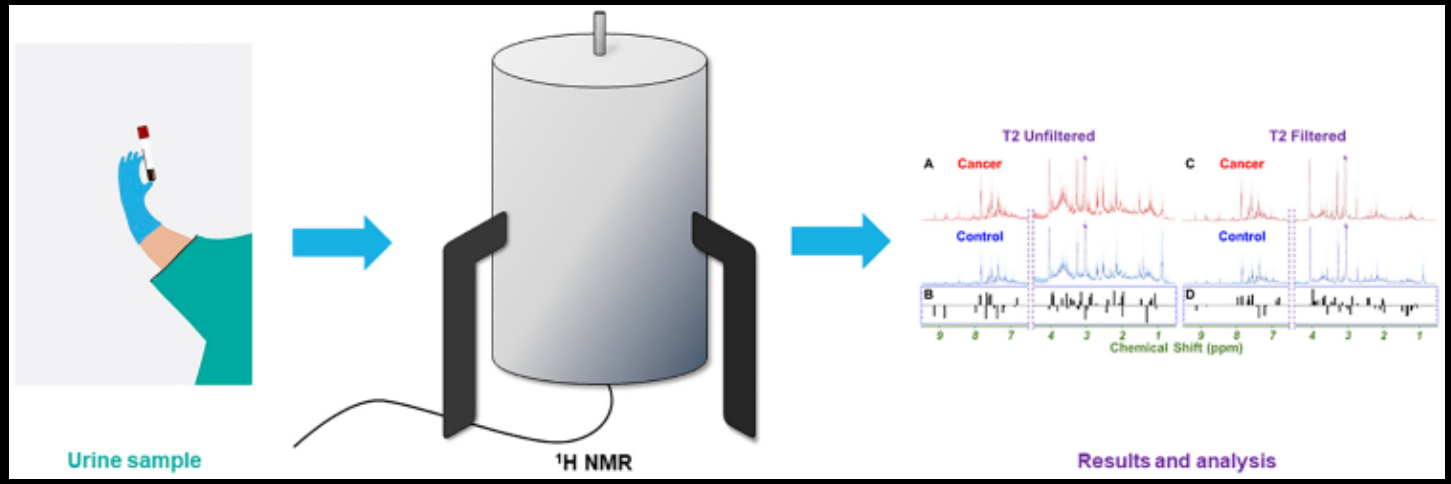
Machine Learning Model for Metabolite Profiling and Quantification in Complex Mixtures from NMR Data

Rabmit Das, Rahul Madhav



Related Works

Bquant

- Database-based identification and quantification of metabolites within local regions of 1H NMR spectra.
- It involves representing spectra as mixtures of reference profiles from a database and inferring the identities and abundances of metabolites through Bayesian model selection.

DeepMID

Deep-learning-based Mixture Identification Method)

- uses a pseudo-Siamese convolutional neural network (pSCNN) and a spatial pyramid pooling (SPP) layer to use to identify plant flavours (mixtures).
- ACC = 99.58%, TPR = 99.48%, FPR = 0.32%; and two experimentally obtained data sets: one shows ACC = 97.60%, TPR = 92.81%, FPR = 0.78%

SMART-Miner

- Tool for metabolite identification from 1H-13C HSQC spectra.
- Based on CNN

	1	5	10	20	30
Top 1	98.8	95.5	92.2	84.6	77.9
Top 2	100	99.3	98.2	94.9	91.6

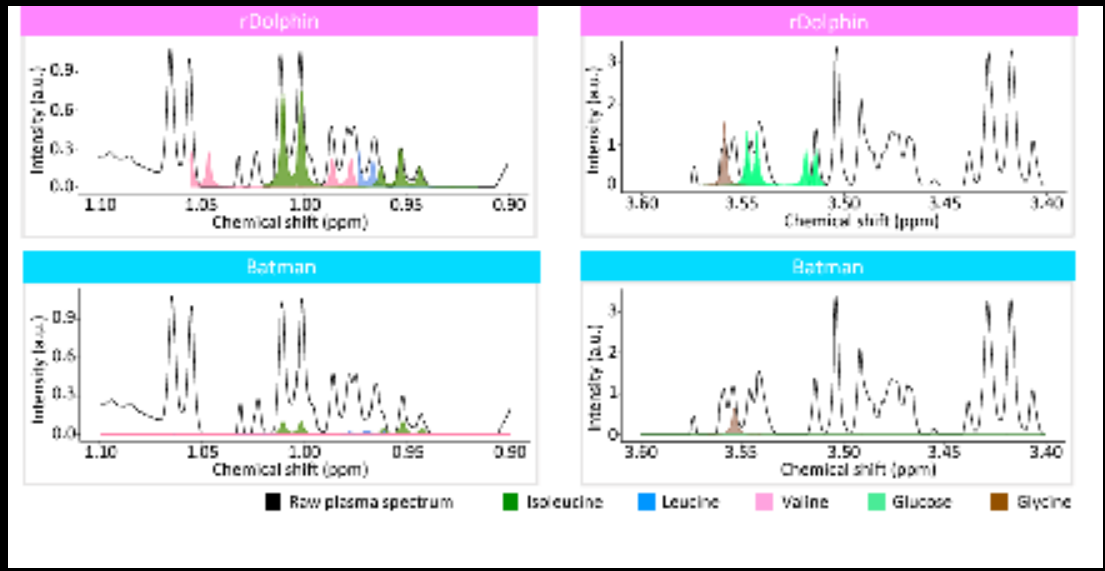
Methodology

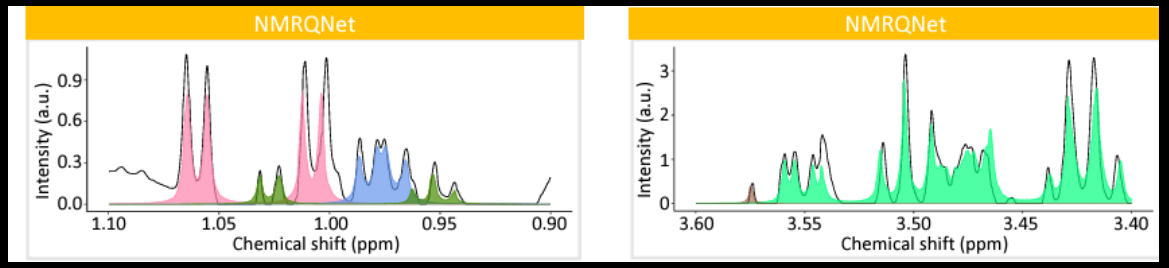
Obtaining NMR of single metabolite → Batman

Data Simulation → Using Rdolphin

Data Preprocessing → PepsNMR

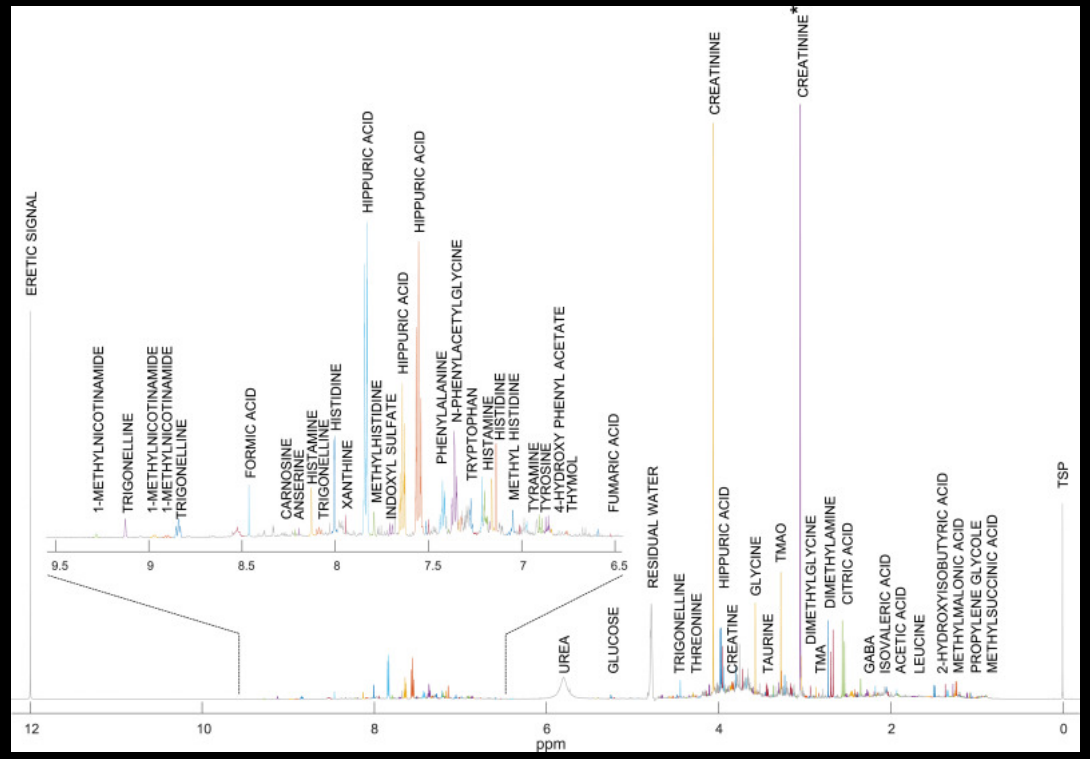
Training Data





Future Targets

1. Generate the synthetic Dataset
2. Run a linear regression model for quantification and a random forest model for profiling.
3. Use two models - First profiling the components of the mixtures (CNN) and second quantifying the classified components (RNN).
4. Combine the two models (CRNN) and compare the individual accuracy.



References

1) Wang, Y., Wei, W., Du, W., Cai, J., Liao, Y., Lü, H., Kong, B.,\ \& Zhang, Z.\ (2023) Deep-Learning-Based mixture identification for nuclear magnetic resonance spectroscopy applied to plant flavors. {\it Molecules, 28(21),}, pp.\ 7380. https://doi.org/10.3390/molecules28217380

2) Zheng, C., Zhang, S., Ragg, S., Raftery, D.,\ \& Vitek, O.\ (2011) Identification and quantification of metabolites in 1H NMR spectra by Bayesian model selection. {\it Bioinformatics, 27(12)}, pp.\ 1637–1644. https://doi.org/10.1093/bioinformatics/btr118

3) Weljie, A. M., Newton, J., Mercier, P., Carlson, E. E.,\ \& Slupsky, C. M.\ (2006) Targeted Profiling: Quantitative analysis of 1H NMR metabolomics data. {\it Analytical Chemistry, 78(13)}, pp.\ 4430–4442. https://doi.org/10.1021/ac060209g

4) https://github.com/danielcanueto/rDolphin

5) https://github.com/LiuzLab/NMRQNet