

BIO-NMR

Machine Learning Approach for Metabolite Profiling in Complex Mixtures via NMR Data Rahul Madhav M, Rabmit Das

Dataset

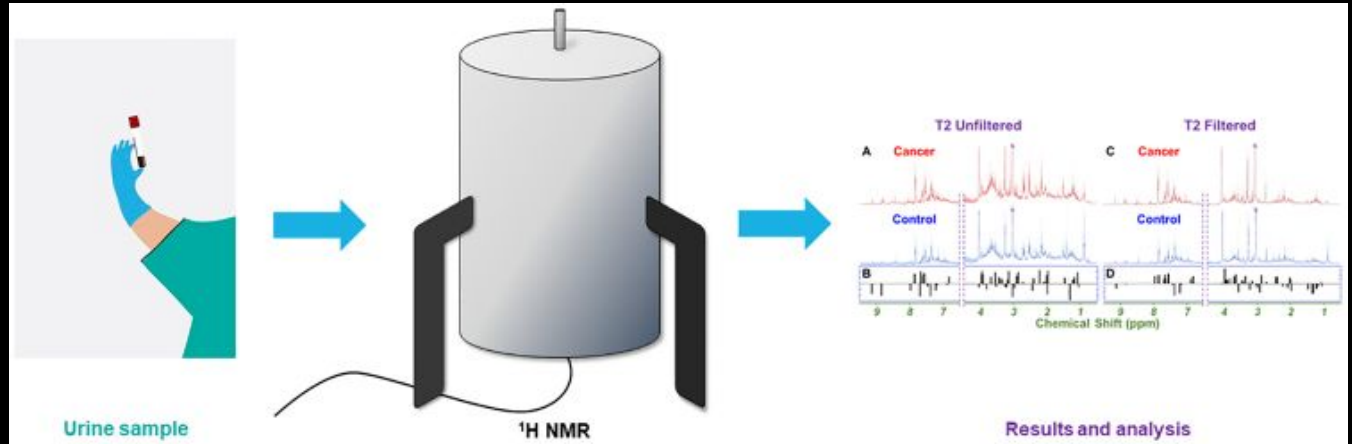
Obtaining NMR
of single
metabolite



Data Simulation



Data
Preprocessing



Relevant papers

1. Schuetzke, J., Szymanski, N. J., & Reischl, M. (2022). A universal synthetic dataset for machine learning on spectroscopic data. arXiv preprint arXiv:2206.06031. <https://doi.org/10.48550/arXiv.2206.06031>.
2. Paul E. Anderson, Michael L. Raymer, Benjamin J. Kelly, Characterization of 1H NMR spectroscopic data and the generation of synthetic validation sets, *Bioinformatics*, Volume 25, Issue 22, November 2009, Pages 2992–3000, <https://doi.org/10.1093/bioinformatics/btp540>.
3. J 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. *Molecules*, 28(21), pp. 7380. <https://doi.org/10.3390/molecules28217380>

Background Idea

We encounter numerous complex mixtures in various fields. Nowadays, NMR spectroscopy has become increasingly prevalent in various industries and research fields. This technique is being widely adopted for the analysis of complex mixtures due to its non-destructive and non-invasive nature. It allows researchers and professionals to simultaneously detect and identify multiple compounds within mixtures without the need for extensive sample preparation or separation steps.

What to do by Midway

1. Literature review.
2. Synthesis of dataset.
3. Implement classic machine learning (Regression) algorithms to achieve the best results.

Post midway work

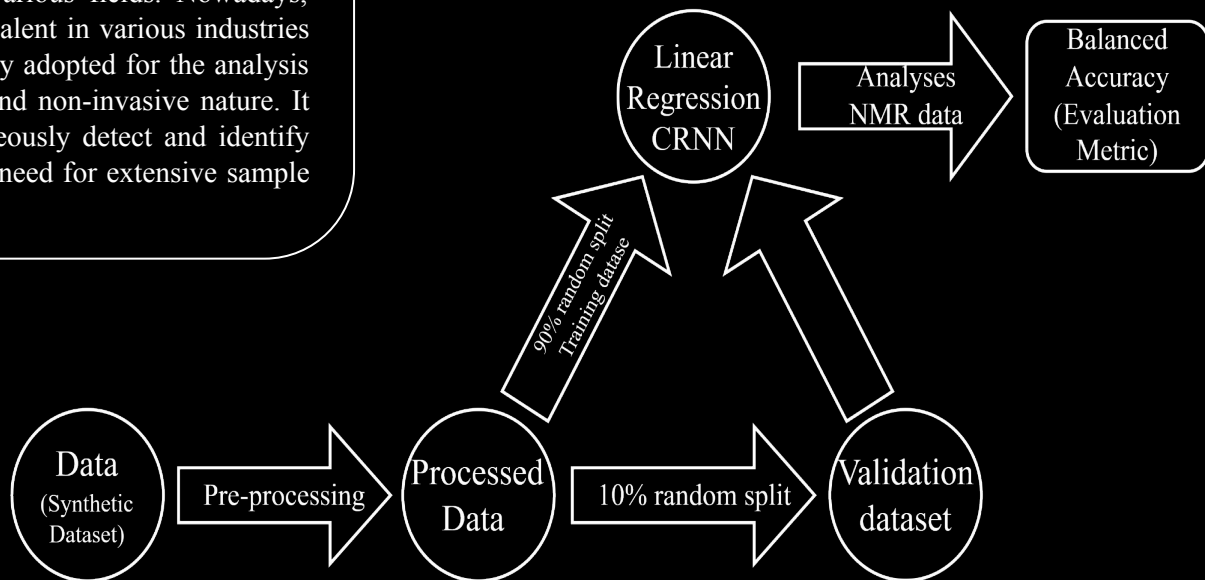
If our primary goals are completed, we will move forward to

1. Learn more about deep learning.
2. Implementation of deep learning model.

Work division

1. Rahul: Data pre-processing.
2. Rabmit: Implementation of different machine learning algorithms.
3. Both: Literature review, Presentation and Report.

Workflow



Expected results

To detect the presence and quantify the composition of complex metabolome.