Rough plan for our BRENDA paper

Teddy Groves

- Methods
- Results
- Discussion

Methods

Methods

1. How we got the data
2. Data processing choices
3. Model specification
4. Implementation and validation of computation

Possible figures

- Workflow graphic of pipeline
- Comparison of raw vs processed data?

Results

Results

Validation

- Graphical posterior predictive checks
- Tests of out-of-sample accuracy

Utility

- Case studies
- App

Possible figures:

- Case study, comparison with and without our method
- Posterior predictive checks
- Workflow graphic showing how results can be used

Discussion

Unique features of our approach

- Full posterior
- Principled approach to predicting parameters with incomplete data
- High coverage
- Results verifiably reflect assumptions in generative model

Our approach vs others

Predict kinetic parameters with chemical signature: Kroll et al. (2021), Li et al. (2021)

- Low coverage due to incomplete data
- Point predictions

Predict kinetic parameters with OLS ANOVA model: Borger, Liebermeister, and Klipp (2006)

- Lack of partial pooling wastes information
- Point predictions

References

Borger, Simon, Wolfram Liebermeister, and Edda Klipp. 2006. “Prediction of Enzyme Kinetic Parameters Based on Statistical Learning.” *Genome Informatics* 17 (1): 80–87.

Kroll, Alexander, Martin K. M. Engqvist, David Heckmann, and Martin J. Lercher. 2021. “Deep Learning Allows Genome-Scale Prediction of Michaelis Constants from Structural Features.” *PLoS Biology* 19 (10): e3001402. https://doi.org/10.1371/journal.pbio.3001402.

Li, Feiran, Le Yuan, Hongzhong Lu, Gang Li, Yu Chen, Martin K. M. Engqvist, Eduard J. Kerkhoven, and Jens Nielsen. 2021. “Deep Learning Based Kcat Prediction Enables Improved Enzyme Constrained Model Reconstruction.” Cold Spring Harbor Laboratory. https://doi.org/10.1101/2021.08.06.455417.