NeurIPS AI4Science 2025 | Dec 7
San Diego, CA

Scaling Multi-Modal and Multi-Task Transformers for Small Molecule Drug Discovery

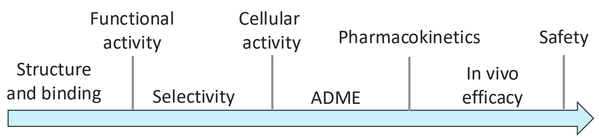
David S. Farina Jr, Sai Krishna Sirumalla, Michiel J.M. Niesen, Daniele A. Di Cesare, Felipe Costas Farias, Michael B. O'Connor, Marcelo Gomes Pereira de Lacerda, Orion Walker Dollar, Peter J. Bygrave, Thomas Dresselhaus, Zhuoran Qiao, Rishi Shah, Jason Swails, Daniel Miles, Oliver Feighan, Stephen Opalenski , Wallace Derricotte, Stephen Munoz, Feizhi Ding, Frederick R. Manby, Thomas F. Miller III, Matthew Welborn

Iambic Therapeutics, Inc, San Diego, CA

Introduction

Motivation

Drug Discovery is about solving many challenges, not one



Multi-task and multi-modal models are needed to maximally utilize drug discovery data

Each stage of the pipeline features diverse types of data, ranging from chemical structures and physicochemical properties to biological assay results and clinical outcomes.

Challenges in applying ML to drug discovery

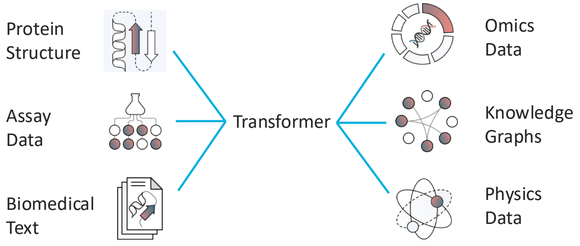
- Nature and quality of available data
- High cost of acquiring high-quality experimental data
- The enormous number of data modalities.

Summary of our approach

Unify and curate disparate large public databases, including PubChem, ChEMBL, and Protein Data Bank.

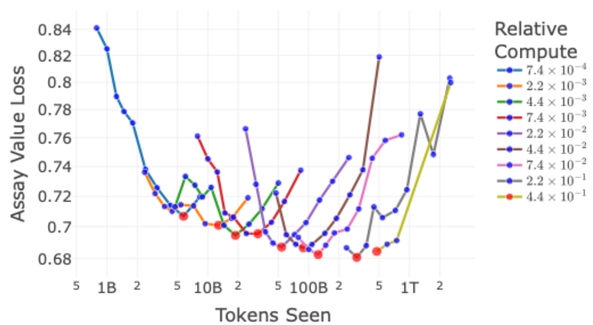
Develop a multimodal, multitask transformer trained on a large, heterogeneous corpus of drug-discovery data. Following the approach introduced with Enchant v1.

Evaluate performance, showing that Enchant v2 performance follows established transformer scaling laws, improving predictably with increased pretraining compute.



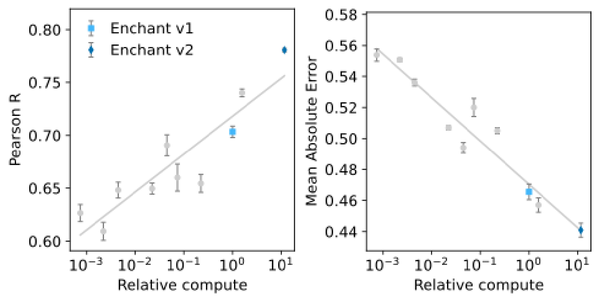
Scaling Laws

Figure 1: Performance (red dot) improves as FLOPs increase



- We performed an isoFLOP scaling study to identify the optimal balance of model size vs. training tokens for Enchant v2.
- (y-axis): Assay value validation loss, cross-entropy loss measuring how well the model predicts measured assay values.

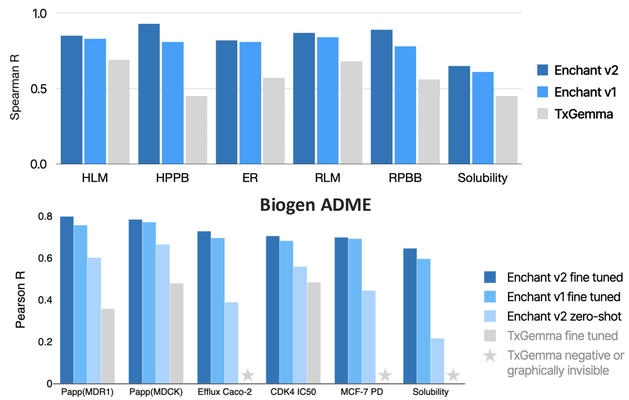
Figure 2: Enchant shows predictable performance improvements through increased model scale



- We fine-tuned each compute-optimal model and observed predictable gains in performance across multiple benchmarks.
- With 10× more pretraining compute than Enchant v1, Enchant v2 achieves substantially better downstream accuracy (Pearson R 0.78 vs. 0.71, MAE 0.43 vs. 0.46).

Benchmarks

Figure 3: Enchant v2 outperforms on Biogen ADME and internal benchmarks

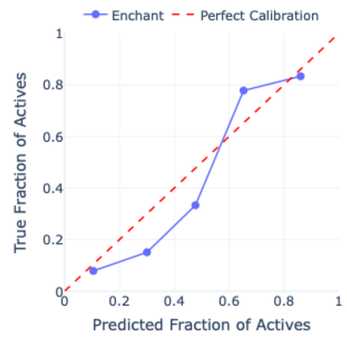


Internal Iambic Assays

- Enchant v2 is fine-tuned for assay prediction using LoRA.
- Figure 3 compares the performance of the Enchant v2 model against Enchant v1 and TxGemma.
- Fine-tuned Enchant v2 achieves state-of-the-art performance.

Applications in Drug Discovery – Hit ID

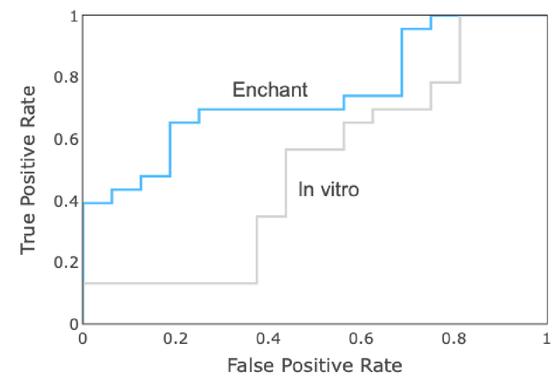
Figure 4: Enchant v2 can improve hit ID



- Compounds were assigned to 5 bins based on predicted pIC50 and bin hit-rate was estimated using model uncertainty.
- For the top 12 predicted compounds, 10 were confirmed hits (83% vs. 86% predicted); while low-ranked compounds showed similarly accurate alignment (8% vs. 10% predicted).

Applications in Drug Discovery – in vivo PK

Figure 5: Enchant v2 can learn from lower-cost in vitro data to predict higher-cost in vivo assay outcomes



- There is a disconnect between in vitro assays and in vivo clearance, limiting compound prioritization.
- Extrapolation from in vitro rat microsomes shows no capacity to predict in vivo clearance (grey line; AUROC 0.51).
- Enchant v2—trained jointly on in vitro and in vivo data—achieves substantially higher accuracy (blue line; AUROC 0.74) in classifying high vs. low clearance compounds.

Conclusion

- Enchant v2 is a large-scale, multimodal transformer for predicting key biochemical, pharmacological, and PK endpoints.
- A unified token representation plus large-scale pretraining yields consistent gains across public benchmarks and internal tasks.
- The model adheres to transformer scaling laws, improving predictably with more compute.
- Enchant v2 improves hit prediction, property estimation, and in vivo PK, including low-data and zero-shot settings.

Work With Us

Internships | Careers | Partnerships