

## Supporting Information

# AweGNN: Auto-parametrized weighted element-specific graph neural networks for molecules

Timothy Szocinski<sup>1</sup>, Duc Duy Nguyen<sup>2</sup>, and Guo-Wei Wei<sup>1,3,4</sup> \*

<sup>1</sup> Department of Mathematics, Michigan State University, MI 48824, USA

<sup>2</sup> Department of Mathematics, University of Kentucky, KY 40506, USA

<sup>3</sup> Department of Biochemistry and Molecular Biology

Michigan State University, MI 48824, USA

<sup>4</sup> Department of Electrical and Computer Engineering

Michigan State University, MI 48824, USA

In the supporting information, we present feature importance heat maps for all 4 toxicity data sets, where we provide 4 separate heat maps for each data set. One heat map shows the sum of the feature importance percentages for each element-specific group derived from the training of the random forest (RF) models on the network-enabled automatic representations (NEARs) generated by the single-task (ST) Auto-parametrized weighted element-specific Graph Neural Networks (AweGNNs). The next heat map displays the feature importances derived from training the RF models on the NEARs generated by the multi-task (MT) AweGNNs. The final two heat maps show the feature importance analogous to the first two, but the feature importance information is derived from training gradient boosting tree (GBT) models. Figures S1-S4 show all the feature importances of each data set obtained from training the RF models on ST NEARs, figures S5-S8 show the feature importances of the RF models trained on the MT NEARs, and the remaining figures, S9-S16, show the feature importances for the GBT models.

---

\*Address correspondences to Guo-Wei Wei. E-mail:wei@math.msu.edu

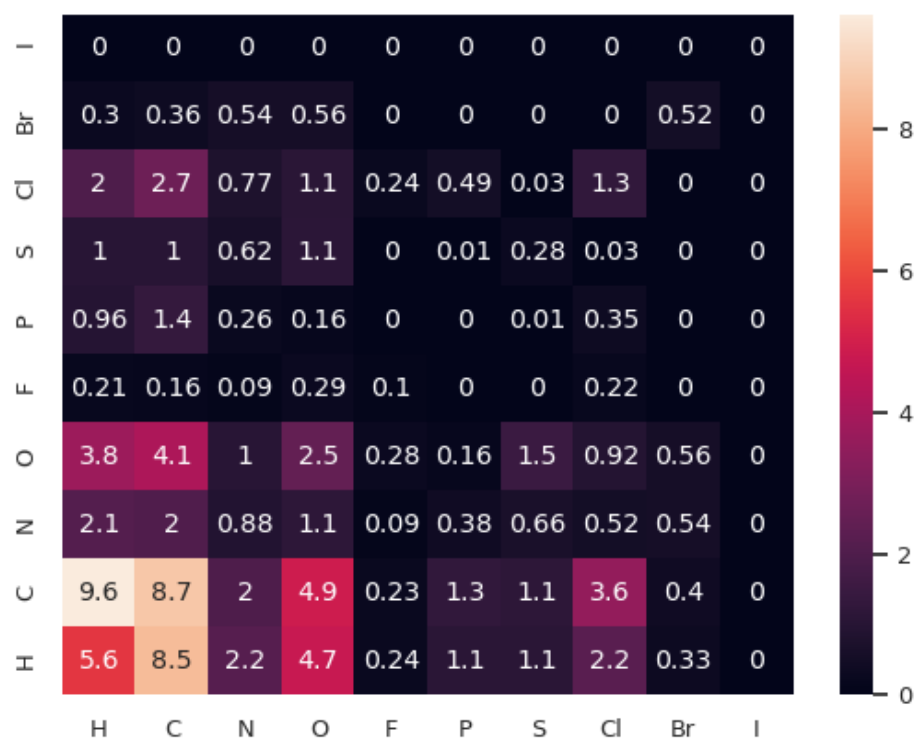


Figure S1: ST LC<sub>50</sub>-DM RF feature importance map.

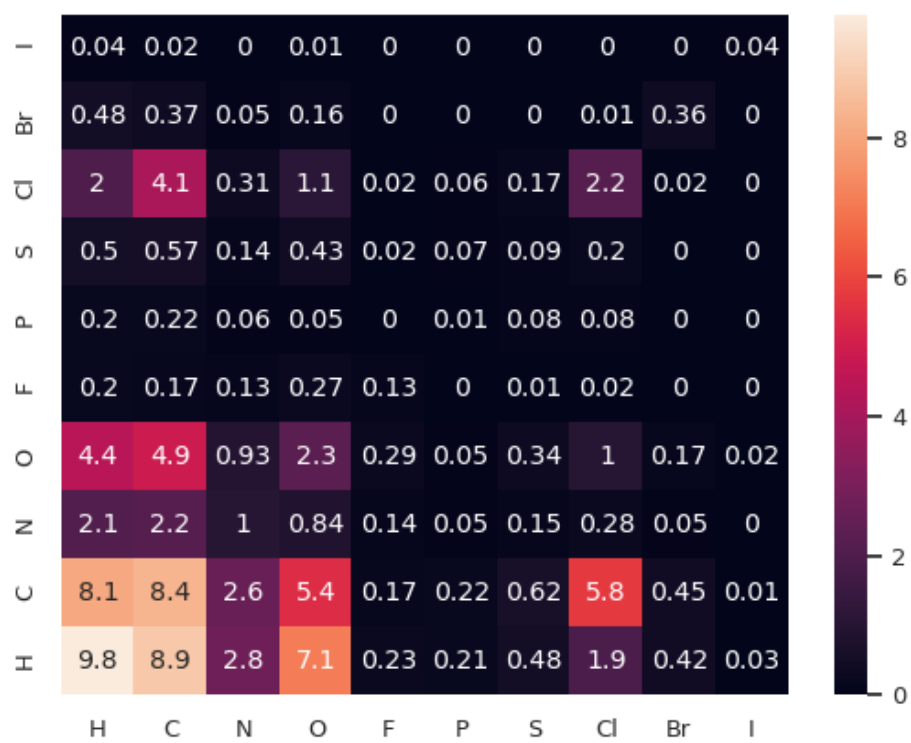


Figure S2: ST LC<sub>50</sub> RF feature importance map.

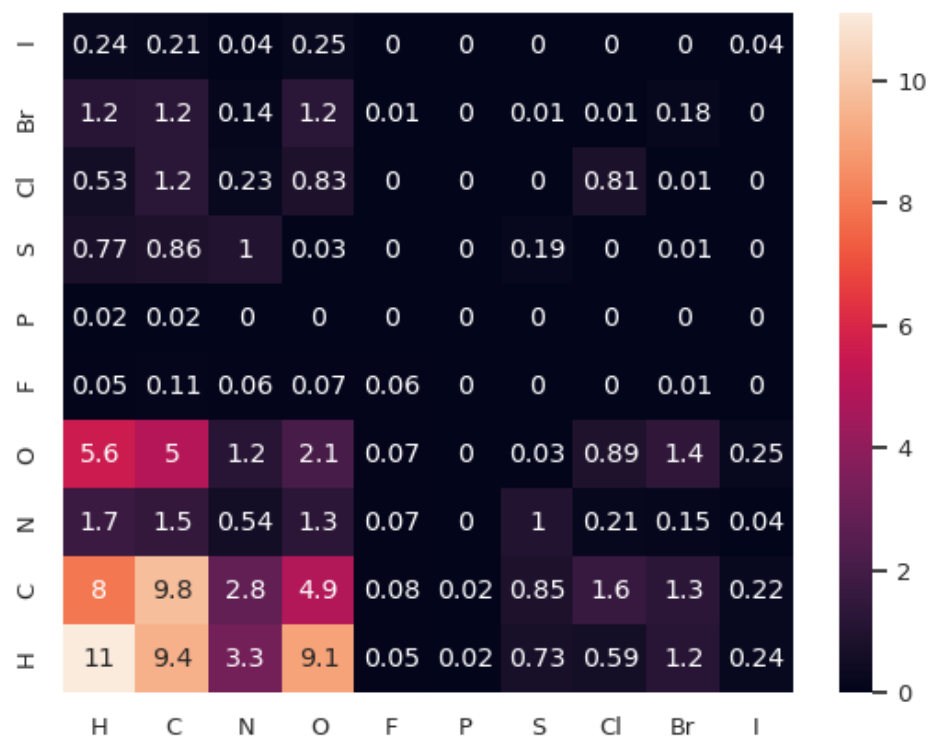


Figure S3: ST IGC<sub>50</sub> RF feature importance map.

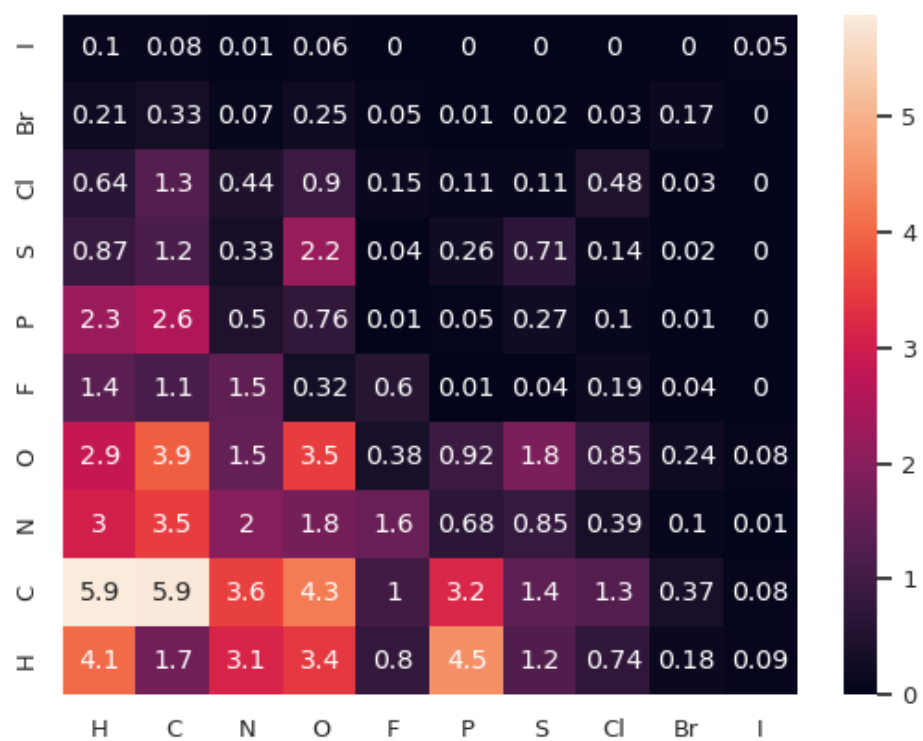


Figure S4: ST LD<sub>50</sub> RF feature importance map.

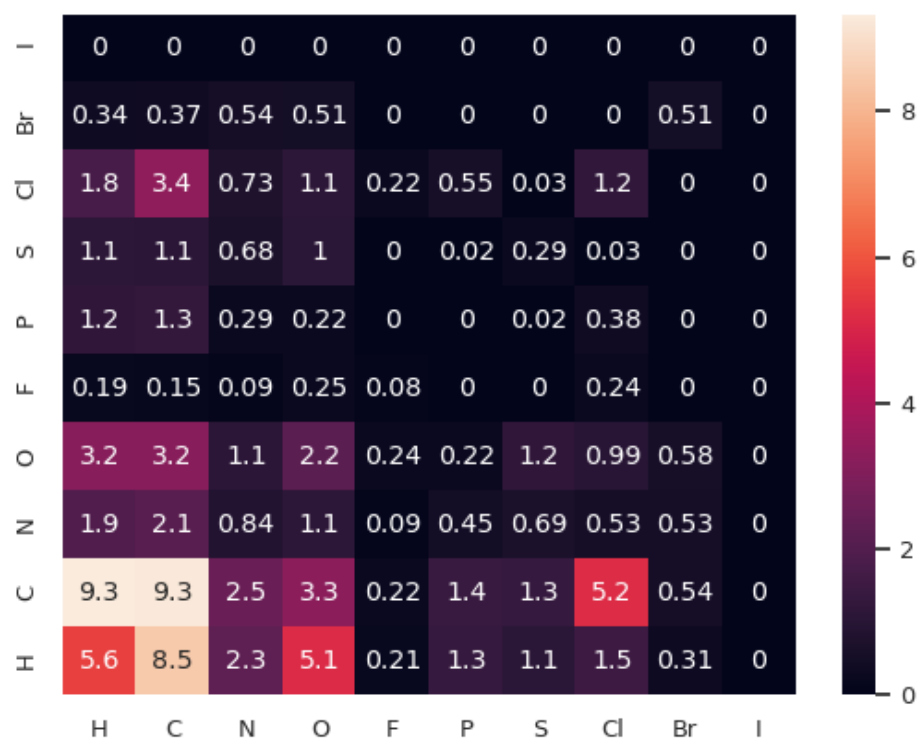


Figure S5: MT LC<sub>50</sub>-DM RF feature importance map.

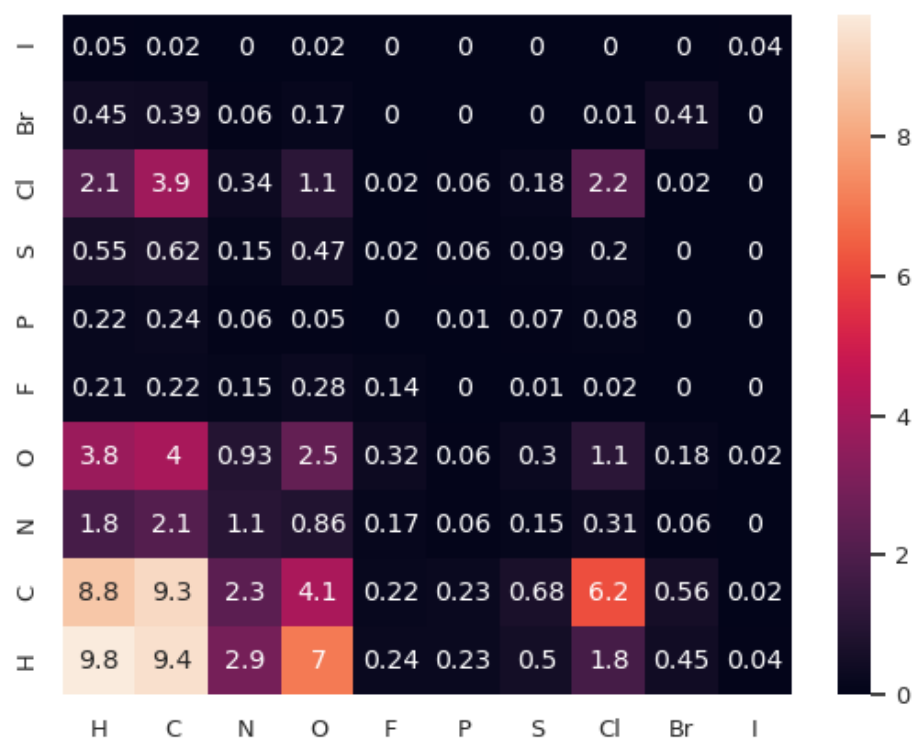


Figure S6: MT LC<sub>50</sub> RF feature importance map.

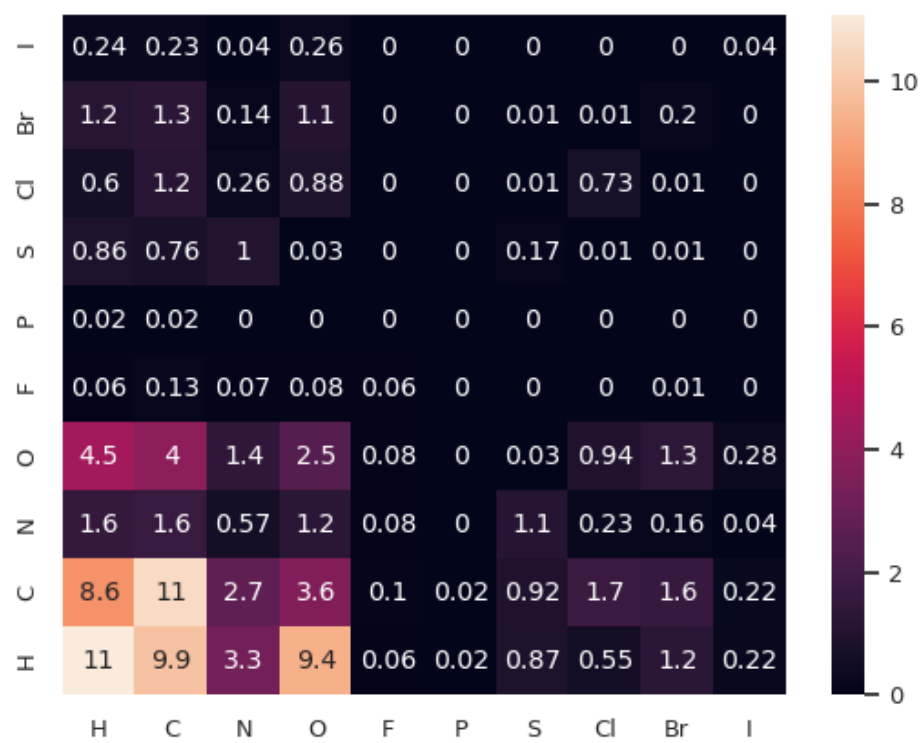


Figure S7: MT IGC<sub>50</sub> RF feature importance map.



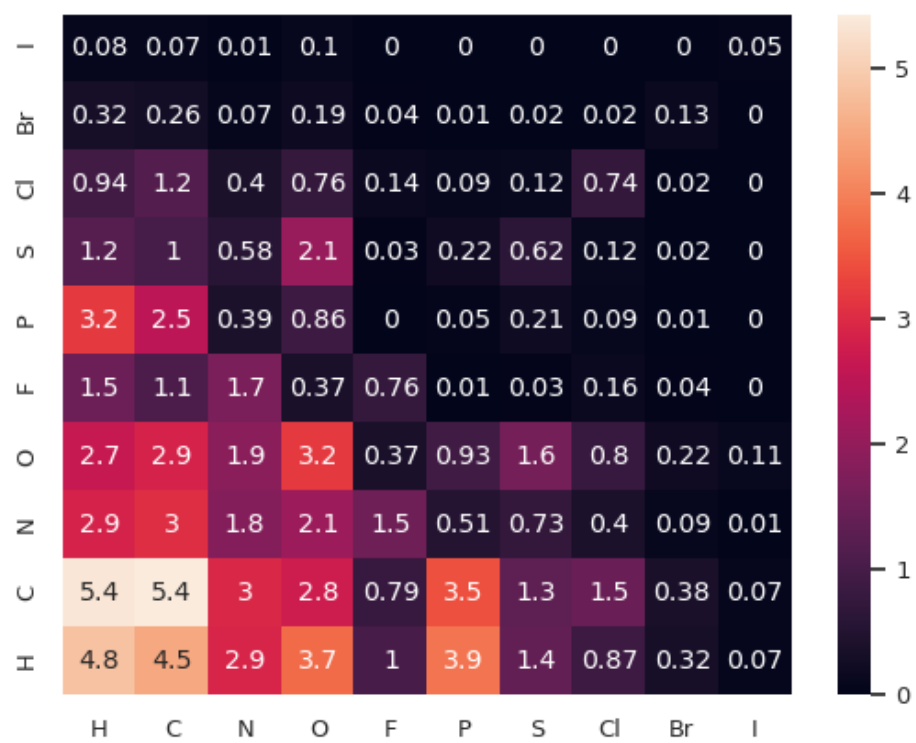


Figure S8: MT LD<sub>50</sub> RF feature importance map.

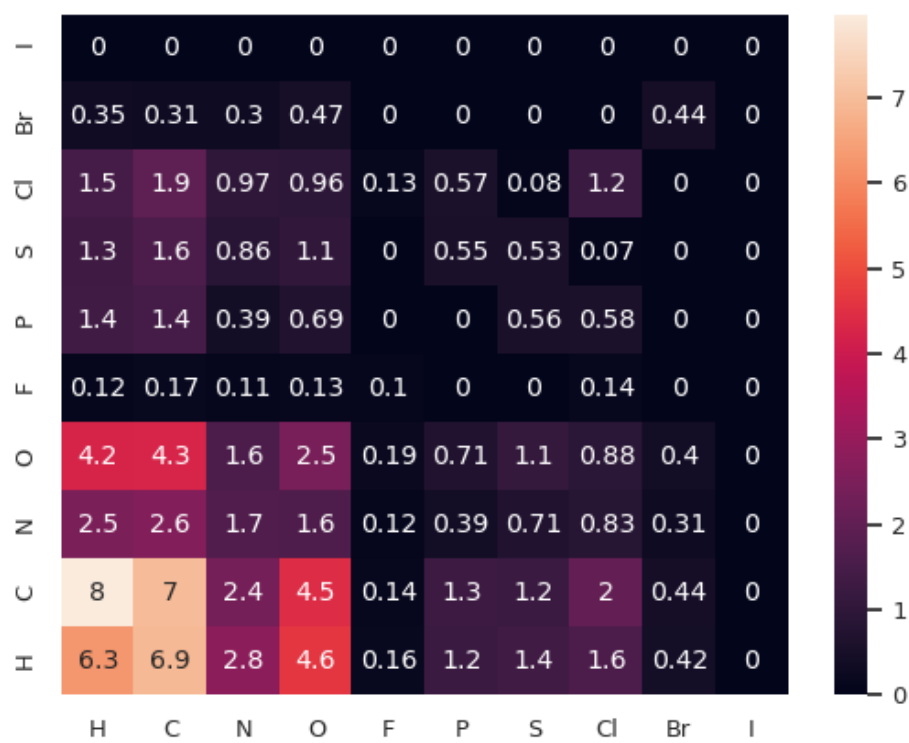


Figure S9: ST LC<sub>50</sub>-DM GBT feature importance map.

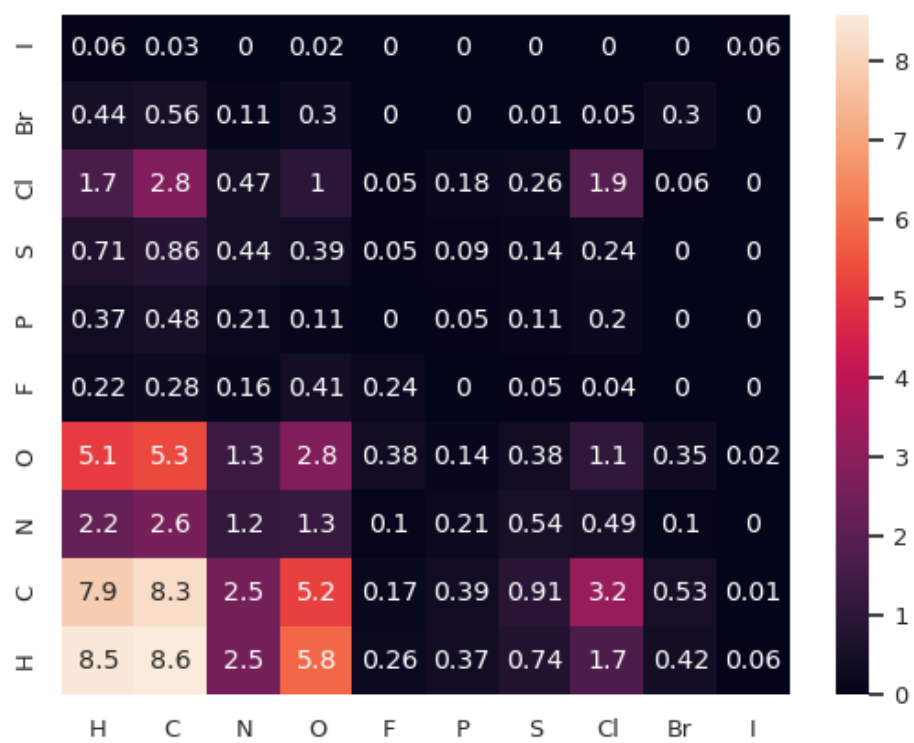


Figure S10: ST LC<sub>50</sub> GBT feature importance map.

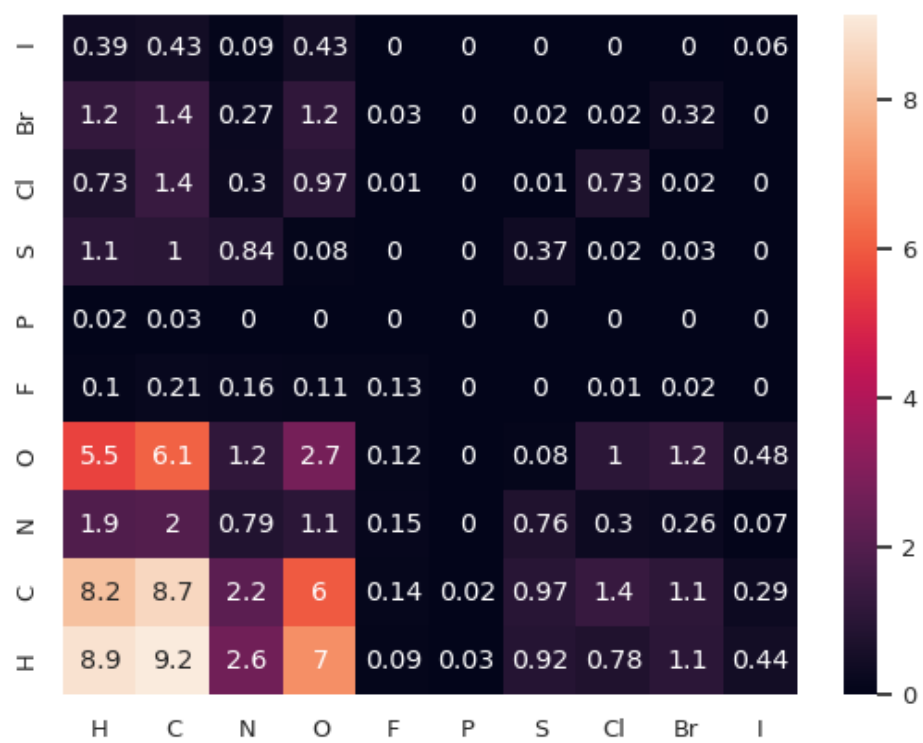


Figure S11: ST IGC<sub>50</sub> GBT feature importance map.



Figure S12: ST LD<sub>50</sub> GBT feature importance map.



Figure S13: MT LC<sub>50</sub>-DM GBT feature importance map.

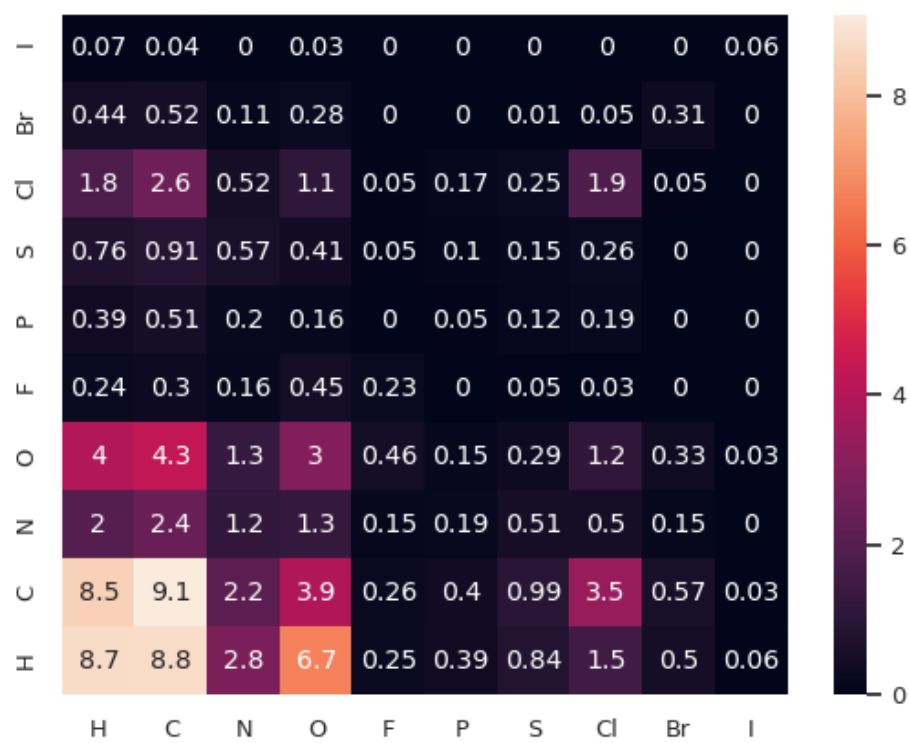


Figure S14: MT LC<sub>50</sub> GBT feature importance map.

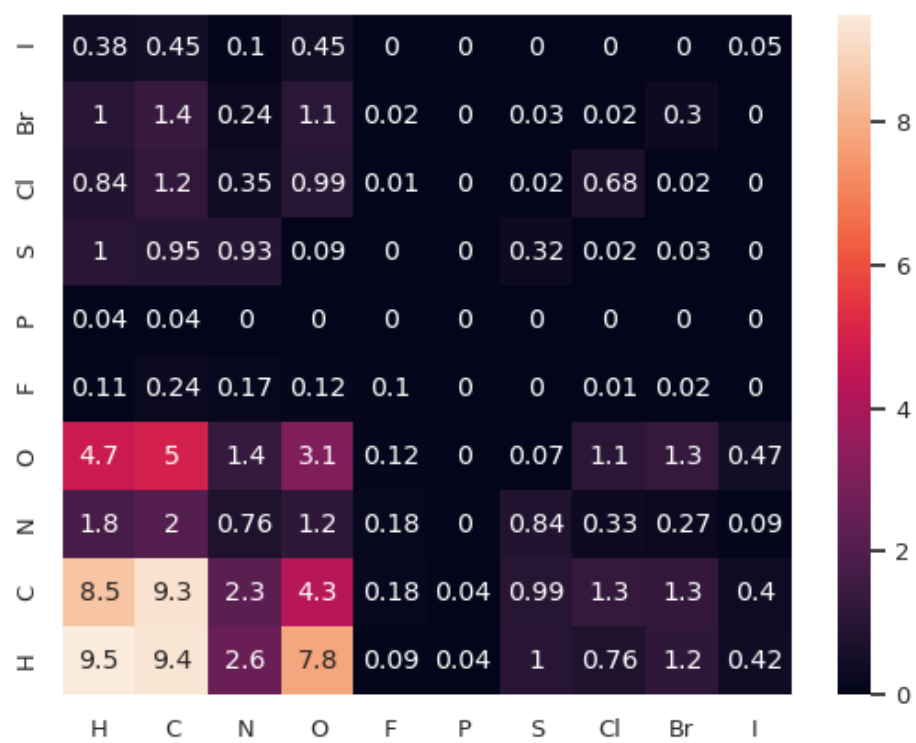


Figure S15: MT IGC<sub>50</sub> GBT feature importance map.



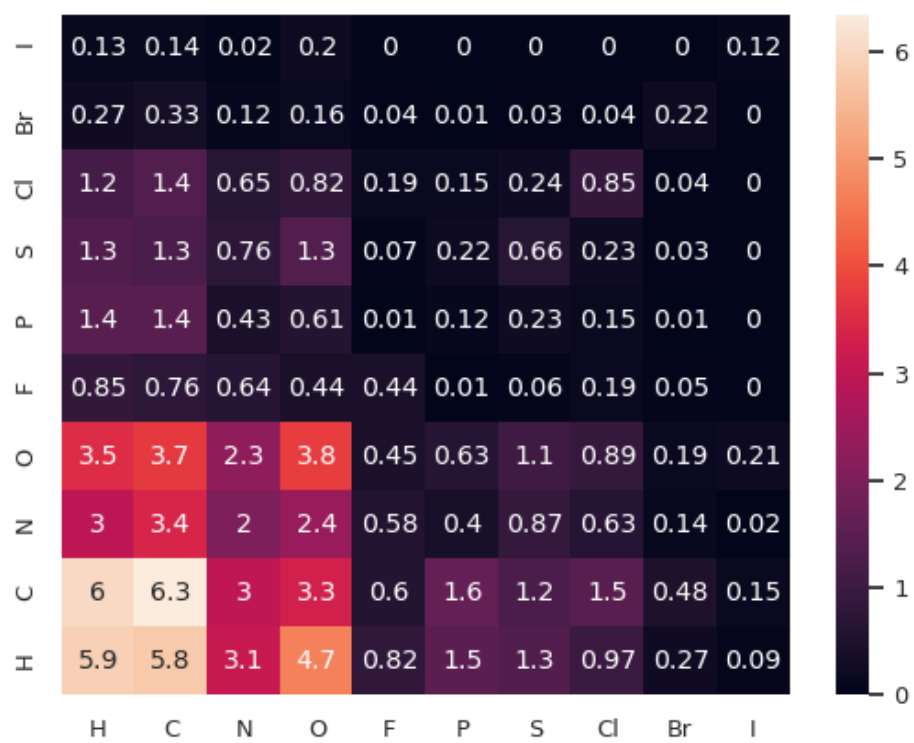


Figure S16: MT LD<sub>50</sub> GBT feature importance map.