

Description of Additional Supplementary Files

1. Supplementary Data 1

This file contains the raw data used to produce the figures within the paper. Specifically, it contains the following datasets in separate sheets:

- i. Fig2a: percentage performance difference between the ACS and MTL-GLC training schemes as a function of task imbalance. The performance is measured in area under the receiver operating characteristic curve (AUC-ROC)
- ii. Fig2b: percentage performance difference between the ACS and STL training schemes as a function of task imbalance. The performance is measured in area under the receiver operating characteristic curve (AUC-ROC)
- iii. Fig2c: percentage performance difference between the MTL-GLC and STL training schemes as a function of task imbalance. The performance is measured in area under the receiver operating characteristic curve (AUC-ROC)
- iv. Fig3: coefficient of determination (R^2) scores of the MTL, MTL-GLC, STL and ACS training schemes over the SAF dataset
- v. Fig4: coefficient of determination (R^2) scores of the MTL, MTL-GLC, STL and ACS training schemes over the SAF dataset segmented into ultra-low data and low data regimes
- vi. Fig6a: latent space representations of the SAF dataset molecules produced by the ACS backbone (i.e., final graph embeddings) for AIT target
- vii. Fig6b: latent space representations of the SAF dataset molecules produced by the ACS backbone (i.e., final graph embeddings) for T_b target
- viii. Fig6c: latent space representations of the SAF dataset molecules produced by the ACS backbone (i.e., final graph embeddings) for LFL target
- ix. Fig6d: latent space representations of the SAF dataset molecules produced by the ACS backbone (i.e., final graph embeddings) for UFL target