**Table 1. Summary of the features used in the classification task.**

|  |  |  |
| --- | --- | --- |
| **Type of feature** | **Name** | **Description** |
| Gene expression features | Landmark genes | 978 genes directly measured from the L1000 datasets |
| DisGeNET DILI genes | Curated genes associated to 9 phenotypes related with DILI from DisGeNET database |
| GUILDify DILI genes | Genes associated through the protein interactions network to 6 phenotypes related with DILI using GUILDify |
| DILI landmark genes | 66 landmark genes selected using non-parametric test for each gene across all samples of Most/Less- vs. No-DILI-Concern drugs (P-value<0.05) |
| Structural features | SMILES | Line notation describing the chemical structure of drugs |
| Drug target genes | Set of targets | 1,664 drug targets retrieved from DGIdb, HitPick and SEA |