Table S. Parameters for optimization in XGB and LGB.

|  |  |  |  |
| --- | --- | --- | --- |
| Parameter | XGBoost | LightGBM | Description |
| max\_depth | int 6 – 11 | - | Maximum depth of a tree. |
| Num\_leaves | - | int 2 – 60 | Maximum number of leaves in a tree. |
| Learning\_rate | float 0 – 1 | float 1e-6 – 1 | Step size shrinkage used in update. |
| Gamma | float 0 – 1 | - | Minimum loss reduction required to make a further partition on a leaf node of the tree. |
| Min\_child\_weight | float 0 – 10 | float 1e-6 – 1 | Minimum sum of instance weight needed in a child. |
| Subsample | float 0 – 1 | float 0.4 – 1 | Subsample ratio of the training instances. |
| Subsample\_freq | - | int 1 – 7 | Subsample frequency in iterations. |
| Colsample\_bytree | float 0 – 1 | float 0.4 – 1 | Subsample ratio of columns. |
| Reg\_alpha | float 0 – 1 | float 1e-6 – 1 | L1 regularization term on weights. |
| Reg\_lambda | float 0 – 1 | float 1e-6 – 1 | L2 regularization term on weights. |

“int” and “float” denote integer and floating-point values. Parameter names are taken from the scikit-learn API of each library.

“learning\_rate”, “min\_child\_weight”, “reg\_alpha” and “reg\_lambda” in LGB were sampled from a logarithmic scale.