Supplementary Table 4 all the prediction models are trained using the optimal hyperparameters

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Hyperparam | Model | | | | |  |
| Model 1 | Model 2 | Model 3 | Model 4 | Model 5 | Model 6 |
| HL0 Node | 2000 | 100 | 1024 | 100 | 100 | 1024 |
| HLi Node | 0.5 | 0.75 | 0.5 | 0.5 | 0.66 | 0.5 |
| Hidden layer | 3 | 3 | 3 | 2 | 5 | 3 |
| Learning rate / eta^ | 0.01 | 0.02 | 0.01 | 0.01 | 0.01 | 0.01 |
| Dropout rate | 0.5 | **0.2** | 0.5 | **0.5** | **0.1** | 0.5 |
| n\_estimators\*^ | - | **-** | - | - | - |  |
| colsample\_bytree^ / max\_samples\* | - | **-** | - | - | - |  |
| max\_depth\* | - | **-** | - | - | - |  |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Hyperparam | Model | | | | |  |  |
| Model 7 | Model 8 | Model 9 | Model 10 | Model 11 | Model 12 | Model 13 |
| HL0 Node | 1024 | 1024 | 1024 | - | - | - | - |
| HLi Node | 0.5 | 0.5 | 0.5 | - | - | - | - |
| Hidden layer | 3 | 3 | 3 | - | - | - | - |
| Learning rate / eta^ | 0.01 | 0.01 | 0.01 | - | 0.5 | - | 0.5 |
| Dropout rate | 0.5 | 0.5 | 0.5 | - | - | - | - |
| n\_estimators\*^ | - | 400 | 400 | 400 | 300 | 400 | 300 |
| colsample\_bytree^ / max\_samples\* | - | 1 | 1 | 1 | 1 | 1 | 1 |
| max\_depth\* | - | 6 | 6 | 6 | - | 6 | - |