**Supplementary Table 3.** The various models (continuous binding strength and the classification of zoonotic capacity or AA30 charge) run for our two datasets (297 vertebrates and 126 mammals) with all associated parameters and evaluation statistics given. The column “Wild/non-wild” refers to how we dealt with domesticated species in our dataset, either by removing them or including a binary variable on the designation of the species. Parameters used in grid search included learning rate (ETA), max depth, and number of minimum observations in each node. For each bootstrap run of 10 iterations, or 50 iterations for the mammal zoonotic capacity models, we recorded the average number of trees, and the training and the test evaluation statistic (AUC or pseudo-R2). Corrected test AUC or pseudo-R2 were calculated using the mean test statistic from an equal number of iterations of a null model.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dataset | Model | Wild/non- wild | Parameters | Evaluation |
| ETA | Max depth | Min obs. in node | Trees | Training eval. | Test eval. | Null test eval. | Corrected eval. |
| Vertebrates | Binding strength | NA | 0.0001 | 4 | 5 | 36,268 | 0.680 | 0.372 | -0.010 | **0.362** |
| Vertebrates | Zoonotic capacity | NA | 0.0001 | 2 | 2 | 59,610 | 0.977 | 0.855 | 0.571 | **0.784** |
| Vertebrates | AA30 | NA | 0.0001 | 2 | 4 | 46,176 | 0.972 | 0.868 | 0.575 | **0.793** |
| Mammals | Binding strength | Removed | 0.0001 | 4 | 3 | 26,889 | 0.754 | 0.157 | -0.048 | **0.109** |
| Mammals | Zoonotic capacity | Removed | 0.0001 | 2 | 5 | 44,652 | 0.997 | 0.843 | 0.618 | **0.725** |
| Mammals | Zoonotic capacity | Variable | 0.0001 | 2 | 5 | 42,093 | 0.993 | 0.789 | 0.589 | **0.699** |
| Mammals | AA30 | Removed  | 0.0001 | 2 | 2 | 29,040 | 1.000 | 0.963 | 0.607 | **0.856** |