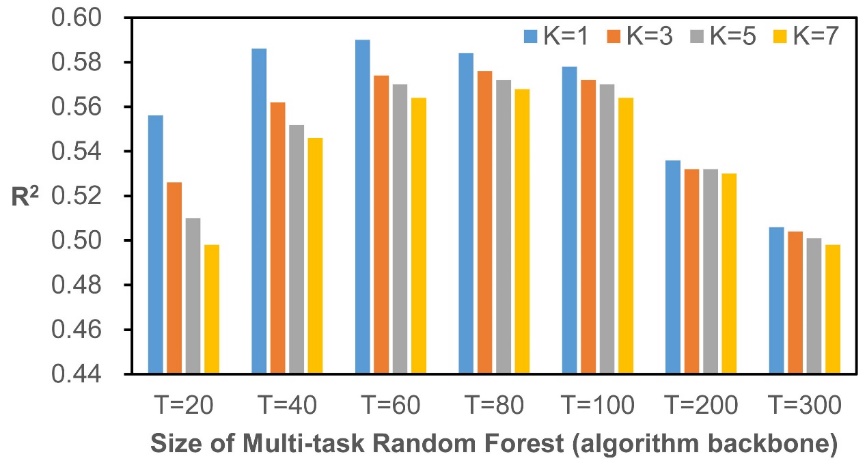
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **SUPPLEMENTARY TABLE I**  **INFORMATION ON THE MULTI-SPECIES ACUTE TOXICITY DATASET USED IN THIS STUDY** | | | | |
| ID | Species | Administration | Toxicity Type | Size |
| 01 | bird-wild | oral | LD50 | 338 |
| 02 | chicken | oral | LD50 | 353 |
| 03 | duck | oral | LD50 | 192 |
| 04 | quail | oral | LD50 | 352 |
| 05 | frog | subcutaneous | LDLo | 112 |
| 06 | mammal (species unspecified) | intraperitoneal | LD50 | 545 |
| 07 | mammal (species unspecified) | oral | LD50 | 674 |
| 08 | mammal (species unspecified) | subcutaneous | LD50 | 125 |
| 09 | mammal (species unspecified) | unreported | LD50 | 1129 |
| 10 | guinea pig | intraperitoneal | LD50 | 248 |
| 11 | guinea pig | intravenous | LD50 | 153 |
| 12 | guinea pig | intravenous | LDLo | 121 |
| 13 | guinea pig | oral | LD50 | 793 |
| 14 | guinea pig | skin | LD50 | 176 |
| 15 | guinea pig | subcutaneous | LD50 | 169 |
| 16 | guinea pig | subcutaneous | LDLo | 179 |
| 17 | mouse | intramuscular | LD50 | 571 |
| 18 | mouse | intraperitoneal | LD50 | 36295 |
| 19 | mouse | intraperitoneal | LDLo | 266 |
| 20 | mouse | intravenous | LD50 | 16978 |
| 21 | mouse | intravenous | LDLo | 102 |
| 22 | mouse | oral | LD50 | 23373 |
| 23 | mouse | oral | LDLo | 264 |
| 24 | mouse | parenteral | LD50 | 302 |
| 25 | mouse | skin | LD50 | 214 |
| 26 | mouse | subcutaneous | LD50 | 6769 |
| 27 | mouse | subcutaneous | LDLo | 252 |
| 28 | mouse | unreported | LD50 | 1739 |
| 29 | rat | intramuscular | LD50 | 300 |
| 30 | rat | intraperitoneal | LD50 | 5021 |
| 31 | rat | intraperitoneal | LDLo | 318 |
| 32 | rat | intravenous | LD50 | 2472 |
| 33 | rat | intravenous | LDLo | 135 |
| 34 | rat | oral | LD50 | 10190 |
| 35 | rat | oral | LDLo | 322 |
| 36 | rat | skin | LD50 | 835 |
| 37 | rat | subcutaneous | LD50 | 1896 |
| 38 | rat | subcutaneous | LDLo | 174 |
| 39 | rat | unreported | LD50 | 806 |
| 40 | cat | intravenous | LD50 | 261 |
| 41 | cat | intravenous | LDLo | 159 |
| 42 | cat | oral | LD50 | 171 |
| 43 | cat | oral | LDLo | 142 |
| 44 | rabbit | intraperitoneal | LD50 | 131 |
| 45 | rabbit | intravenous | LD50 | 792 |
| 46 | rabbit | intravenous | LDLo | 346 |
| 47 | rabbit | oral | LD50 | 894 |
| 48 | rabbit | oral | LDLo | 249 |
| 49 | rabbit | skin | LD50 | 1495 |
| 50 | rabbit | skin | LDLo | 181 |
| 51 | rabbit | subcutaneous | LD50 | 156 |
| 52 | rabbit | subcutaneous | LDLo | 241 |
| 53 | dog | intravenous | LD50 | 468 |
| 54 | dog | intravenous | LDLo | 360 |
| 55 | dog | oral | LD50 | 649 |
| 56 | dog | oral | LDLo | 187 |
| 57 | human | oral | TDLo | 140 |
| 58 | women | oral | TDLo | 156 |
| 59 | man | oral | TDLo | 163 |
| **Total sample size** | | | | **122594** |

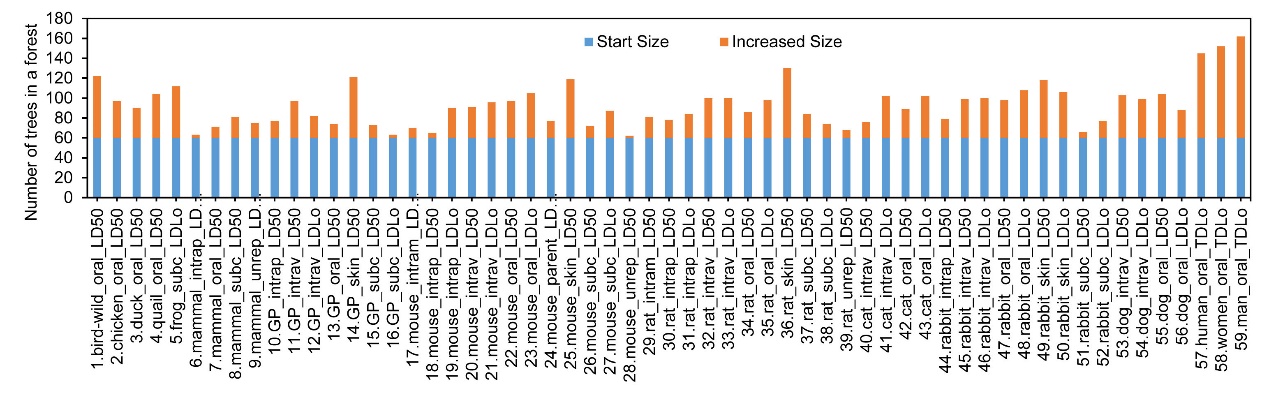
## G. Parameter Analysis of MTIEC-RF

The training of MTIEC-RF involved a structured three-step process: (1) creation of the MT-RF backbone using the training set including 59 endpoints; (2) iterative generation of a set of ECDTs based on the hard sample set (challenging samples) in each endpoint; and (3) application of a composite approach, combining the MT-RF backbone and the corresponding ECDTs, to predict a specific endpoint. Notably, the number of ECDTs was dynamically determined through a pragmatic greedy strategy, and the size of the base classifiers of the final model equaled the sum of the sizes of the MT-RF and ECDTs.

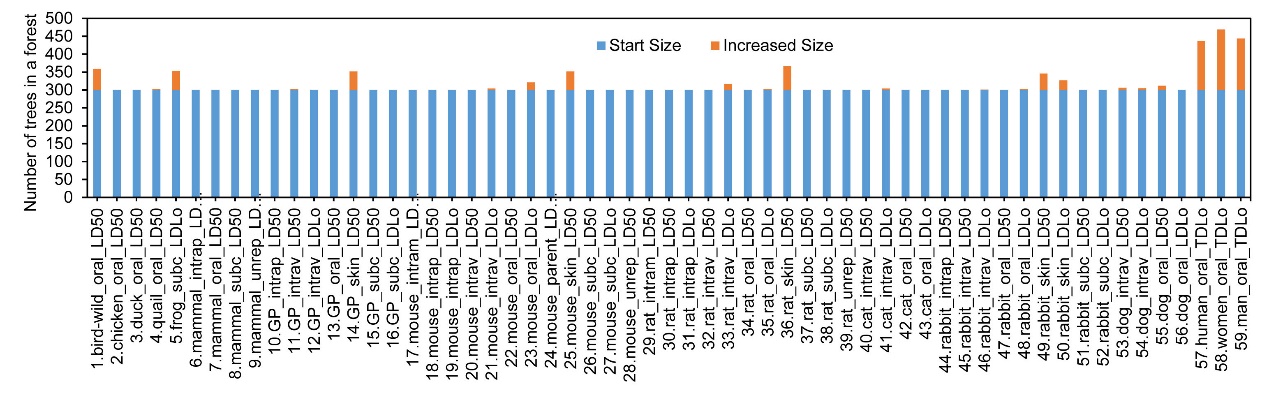
The algorithm included two pivotal hyperparameters: the size of the MT-RF in step 1 (designated as parameter T) and the number of hard samples and neighbors in step 2 (termed parameter K). The impact of parameters T and K on the model performance is depicted in Fig. 6. Evidently, the optimal performance was achieved with T=60 and K=1, aligned with the default parameters of the proposed algorithm. However, as K remained at 1 and T increased from 60 to 300, the performance of the model gradually decreased. Therefore, opposite to the general knowledge that a larger ensemble size promises better performance, our backbone doesn’t require a large number of trees. Indeed, with too many base learners, the overfitting problem tends to occur due to the difference of data distribution across various tasks.



**Fig. 6. Influence of parameter settings on model performance.** The X-axis represents T increasing from 20 to 300, the Y-axis represents the R2 value, and the heights of the columns in different colors represent the model performance at the corresponding K and T values.



**Fig. 7. Growth of ECDTs when the size of MT-RF is 60.**



**Fig. 8. Growth of ECDTs when the size of the MT-RF is 300.** The X-axis represents 59 endpoints, the Y-axis represents the final size of the model, the blue part of the column represents T, the yellow part represents the growth of ECDTs, and the total height of the column represents the final size of the model generated by the proposed algorithm at each endpoint.

Fig.s 7 and 8 illustrate the growth of the ECDTs in the proposed algorithm across each toxicity endpoint for T values of 60 and 300, respectively. As shown in Fig. 7, with setting T to 60, different numbers of ECDTs were generated across different endpoints. Fig. 8 shows that when T was increased to 300, the increase in the number of ECDT remained modest. This result could be caused by the fact that, with an excessively large MT-RF size, the initial addition of an ECDT could not enhance model performance, leading to algorithm convergence. The incorporation of a performance evaluation mechanism, such as an early stop strategy [23], can potentially mitigate premature algorithmic convergence.