

1. Consider an optimized RF model (M) built on the set of autosomal chromosomes  $\{k|i \notin k\}$  binned at some resolution  $r$
2. **for each chr  $i$  do**
  3. Construct the base-level resolution predictor space  $A_{n \times p}$  where  $n$  is the length of chr  $i$  and  $p$  is the number of predictors
  4. Assign threshold  $\{t|0 \leq t \leq 1\}$  and  $\{\epsilon|\epsilon > 0\}$
  5. **if  $|t| > 1$  or  $|\epsilon| > 1$  then**
    6. Assign grid ( $g$ ) of combinations of  $t$  and  $\epsilon$
    7. **for each combination  $l$  in  $g$  do**
      8. Evaluate M on  $A_{n \times p}$  to get the probability of each genomic coordinate as being a domain boundary  $\pi_n$
      9. Subset  $\{\pi_n|\pi_n \geq t_l\}$
      10. Construct the pairwise distance matrix  $D$  between genomic coordinates where  $\pi_n \geq t_l$
      11. Apply DBSCAN on  $D$  with  $MinPts = 3$  and  $\epsilon = \epsilon_l$
      12. **for each cluster  $k$  identified by DBSCAN do**
        13. Assign  $w_k$  as the number of coordinates that span each cluster of bases in  $k$  (PTBR)
        14. Perform PAM on the sub-distance matrix  $D_k$  to extract the cluster medoid  $b_k$  (PTBP)
        15. **for each predictor  $p$  do**
          16. Calculate the normalized enrichment (NE) over all predictors
$$NE = \frac{1}{p} \left[ \sum_{s=1}^p \left[ \frac{1}{b} \sum_{k=1}^b e_{ks} \right] \right]$$

where  $e_{ks} = \mathbf{I}\{r_s \in (b_k - f, b_k + f)\}$  is the number of elemental regions  $r$  of predictor  $p$  that overlap with each flanked boundary
          17. Determine where  $NE$  converges as optimal  $\{t, \epsilon\}$  combination
          18. Repeat steps 12-18 on  $A_{n \times p}$  with optimal  $\{t, \epsilon\}$
      19. **end**
    20. **end**
  21. **end**

**Algorithm 1:** Psuedocode for *preciseTAD* implementation.