Pseudo-code of the model Tree Induction Algorithm (TRIAL)

FUNCTION: TRIAL

Initialise ModelTree structure

Flag root node as open node

WHILE GrowTree = true
    Find open leaf nodes
    
    FOR EACH open leaf node
        Find best split for open leaf node
        Flag node as closed node
    ENDFOR

    Calculate new tree BIC for each leaf node split

    Find best node to split: min(NewTreeBIC)

    IF NewTreeBIC < CurrentTreeBIC
        Split node
        Flag new children leaf nodes as open nodes
    ENDIF ELSE
    GrowTree = false
ENDELSE
ENDWHILE

RETURN: ModelTree

END
Pseudo-code of the function that determines the best split for continuous variables

FUNCTION FindBestContinuousVariableSplit

FOR EACH continuous split variable
    FOR EACH split location
        LeftMembers= WHERE(CurrentSplitVariable < CurrentValue)
        RightMembers= WHERE(CurrentSplitVariable >= CurrentValue)
        IF (nsamplesLeft >= MinSamples) AND (nsamplesRight >= MinSamples)
            Compute multiple regression for left and right child
            Compute the sum of left and right SSE
        ENDIF
    ENDFOR
ENDFOR

Find best variable and location: min(SSE)

Estimate coefficients for left and right children using stepwise variable forward selection; selection criterion: significance and reduction of BIC

Estimate unbiased sum of squared errors for left and right child using crossvalidation

RETURN: Best continuous split, coefficients, crossvalidated errors for left and right child

END
Pseudo-code of the function that determines the best split for categorical variables

FUNCTION: FindBestCategorialSplit

FOR EACH categorical variable

    REPEAT UNTIL two categories are left
    
    Find best combination of two categories in sharing the same multiple regression: \( \min(\text{SSE}) \)

    Merge these two categories

    ENDREPEAT

ENDFOR

Select the variable where splitting results in largest error reduction: \( \min(\text{SSE}) \)

Estimate coefficients for left and right children using stepwise variable forward selection; selection criterion: significance and reduction of BIC

Estimate unbiased sum of squared errors for left and right child using crossvalidation

RETURN: Best categorical split, coefficients, and crossvalidated errors for left and right child

END
Pseudo-code of the model tree ensemble method (Evolving tRees with RandOm gRowth ERROR)

PROCEDURE: ERROR

Grow deterministic tree with TRIAL

REPEAT UNTIL forest complete
  Choose an existing tree that will be modified: Min(random * TreeBIC_rank)
    load tree from working directory
    Choose a random interior node with more than two leaf nodes that will be Pruned
    prune tree at this node
    grow tree randomly from this node
    grow tree deterministically from the new leaf nodes
    IF not all final split nodes are deterministic
      prune tree at final random split nodes
      make deterministic splits at final split nodes
    ENDIF
    save tree in working directory and store the BIC of the tree
ENDREPEAT

Select diverse trees for the ensemble

END