Algorithm: Fast data collection for MAP regression (single layer type).

Input: pre-trained model Y_0, number of rounds rds, pruning budget k, total layers L

Initialize $train\_data = \{(\tilde{m}_{a,0}, \tilde{m}_{g,0}, a_0)\}$;

Function `PruneAlong(`*layer-type*, x_0, $x_{\max}$`):`

  for $n = 1$ **to** rds **do**

    $x_n = x_0 + n \cdot \frac{x_{\max} - x_0}{rds}$;

    Prune Y_{n-1} along *layer-type* to ratio x_n to obtain Y_n;

    Fine-tune Y_n and evaluate $\rightarrow (\tilde{m}_{a,n}, \tilde{m}_{g,n}, a_n)$;

    Append $(\tilde{m}_{a,n}, \tilde{m}_{g,n}, a_n)$ to $train\_data$;

  end

Set $\tilde{m}_{a,\max} = k/L$ and $\tilde{m}_{g,\max} = k/L$;

`PruneAlong(`*"attention"*, $\tilde{m}_{a,0}$, $\tilde{m}_{a,\max}$`);`

`PruneAlong(`*"activation"*, $\tilde{m}_{g,0}$, $\tilde{m}_{g,\max}$`);`

return $train\_data$