**Algorithm**

1. Load dataset A, B, and C into memory.
2. For each dataset:
3. Split the data into training and testing sets.
4. Implement the Results Tree Algorithm, Random Forests, and XGBoost models on the training set.
5. Extract the feature importance or ranking scores of each input variable from the trained model.
6. Sort the feature importance scores in descending order.
7. Select the top K features with the highest importance scores, where K is a predetermined value or based on a threshold.
8. Record the selected features in a table or log.
9. Repeat steps b-f for each dataset.
10. Compare the important features across the different datasets to identify any common or unique features.
11. If the important features are the same for all datasets, select the data from the data set that has the highest number of important features and load to P dataset.
12. If the important features are different for the datasets, then :
13. If the important features are sourced from highest number of important features dataset, then combine the features and data into data set P.
14. Else, use a linear algorithm to predict missing values and create the data for the important features and load into data set P.
15. Repeat steps a-b for each features.
16. Combine the predicted data and features into data set P to ensure that the created data is relevant.

**Example of output**

