Plain Language Summary: This study developed a machine learning technique to retrieve air temperature profile using hyperspectral infrared sounding data observed by a geostationary satellite with two typhoon cases. This study adopted generalized ensemble learning approach to combine three basic models: Random Forest, XGBoost, and LightGBM. The retrieved results were validated with the radiosondes and reanalysis data. As a result, the retrieval accuracy was improved enough, compared with previous studies.
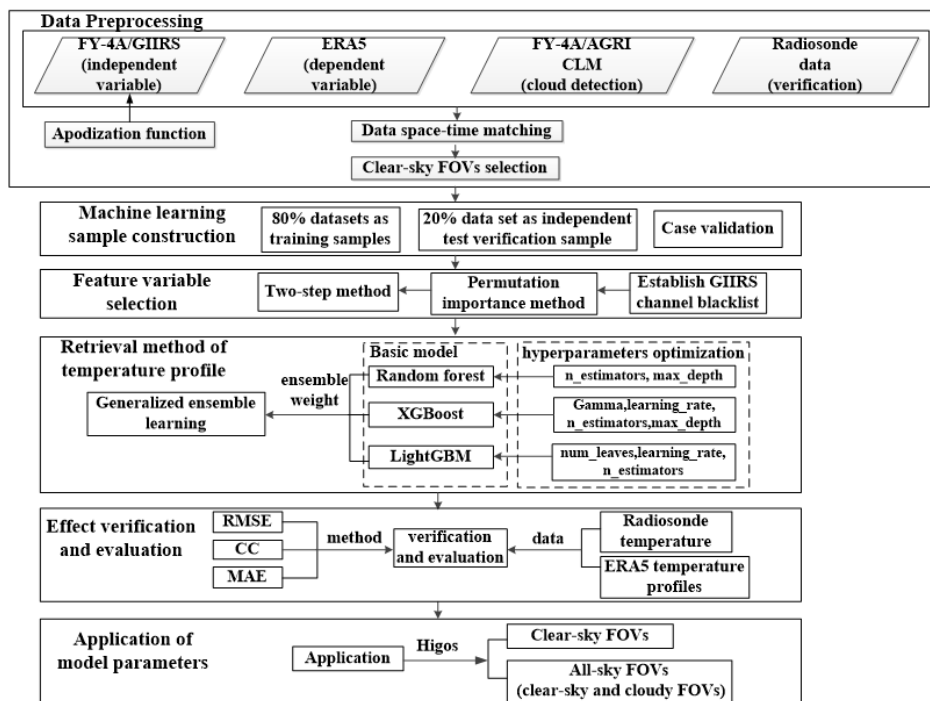

Figure 2. Logical relation framework and flow chart of the generalized ensemble learning retrieval of the temperature profile.

Highlights:

- A two-step feature selection method is used to select important feature variables for the FY-4A/GIIRS medium wave channel.
- Generalized ensemble learning nonlinear convex optimization is used to optimize the weight of each basic model (Random Forest, XGBoost and LightGBM).
- Under clear-sky FOVs, the RMSE of the temperature profile of generalized ensemble learning in the training dataset is less than 0.3 K, and the RMSE of the test dataset is less than 1 K between 150 hPa and 925 hPa.
- Preliminary exploration of temperature retrieval under all-sky (clear-sky and cloudy sky) FOVs also achieves high retrieval accuracy.