## Inclusive Multiple Modeling (IMM) for Calculating SVI

Figure 1 depicts the IMM formulation for calculating SVI at two levels, in which Level 1 ALPRIFT data layers are input datasets, and the normalized subsidence by InSAR processing is the label dataset. Also, at Level 2, the predicted SVI by Level 1 models is input datasets for Level 2 model, and the label dataset is the same as Level 1. In this formulation, the RF and SVM are the Level 1 models, and RF plays the combiner role at Level 2.

A diagram of a forest model

Description automatically generated

**Figure 1.** Mehtodological flowchart for the SVI prediction at two levels

Random Forests (RF) is a tree-structured ensemble learning technique, extended in many fields for regression and clustering due to its performance and simple structure (Breiman 2001; Lee et al. 2024). Two random features of RF improve its performance, which include the aggregation with repeated placement or bagging and the random selection of features. The basic idea of bagging is to generate multiple training sets using bootstrap random sampling from the original training set. Then, using these training sets, different models are trained, and their results are aggregated to reach the final decision. One of the classic applications of bagging in RF is that this bagging method is combined with the decision tree method. However, the decision tree training process in RFs has changed, wherein each node of the tree and the features are randomly selected, and the branch-cutting process is not implemented. Due to these two features, RF reduces variance and deviation (Gordon et al. 1984).

References:

Breiman, Leo. 2001. “Documentation for R Package RandomForest.” Machine Learning.

Gordon, A. D., L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone. 1984. “Classification and Regression Trees.” Biometrics. doi: 10.2307/2530946.

Lee, H., Jeong, J., & Choung, S. (2024). Improvement of performance of in-situ virtual monitoring system of the occurrence probability for high concentrations of naturally occurring radioactive materials in groundwater through the solution of the data imbalance problem. Environmental Modelling & Software, 105978.