Stacking

feature1	feature2	feature3	prediction

Assuming that we have the above training dataset (e.g., 14 samples, 3 features)

Step 1: Split training data to n folds

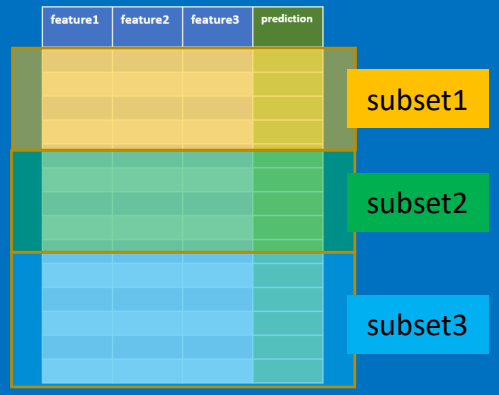
feature1	feature2	feature3	prediction

Assuming that we have the above training dataset (e.g., 14 samples, 3 features)

Step 1: Split training data to k folds

In this case let's assume k=3, so the training data will be split to 3 subset

feature1	feature2	feature3	prediction

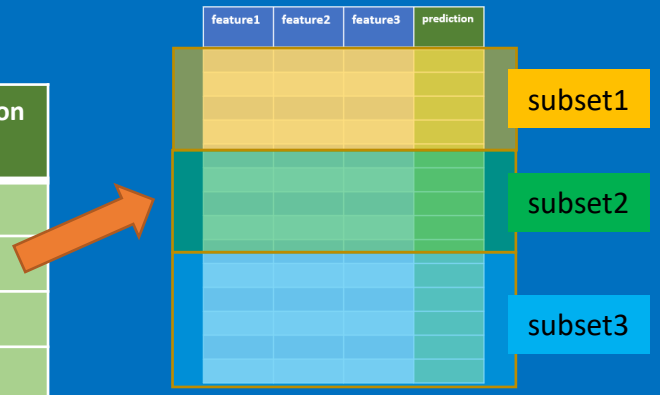


Assuming that we have the above training dataset (e.g., 14 samples, 3 features)

Step 1: Split training data to k folds

In this case let's assume k=3, so the training data will be split to 3 subset

feature1	feature2	feature3	prediction



Fold 1:

- training data: subset1 subset2
- Testing data: subset3

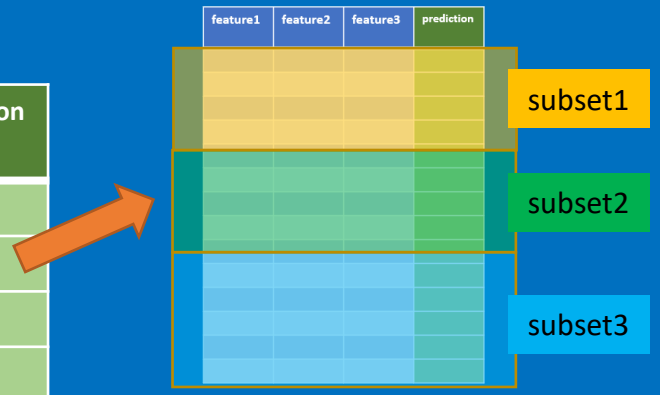
We can use 2 subsets for training and 1 subset for testing

Assuming that we have the above training dataset (e.g., 14 samples, 3 features)

Step 1: Split training data to k folds

In this case let's assume k=3, so the training data will be split to 3 subset

feature1	feature2	feature3	prediction



Fold 1:

- training data: subset1 subset2
- Testing data: subset3

Fold 2:

- training data: subset1 subset3
- Testing data: subset2

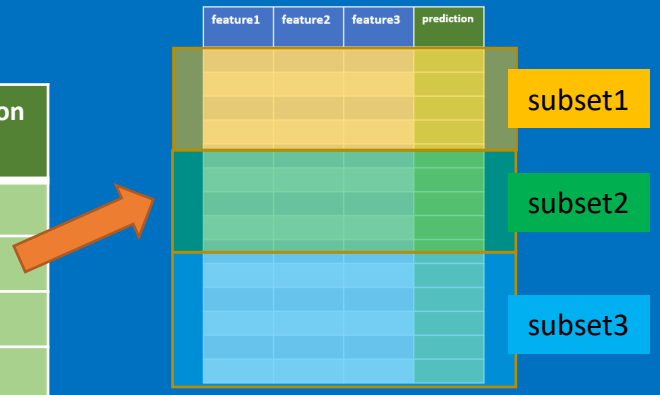
We can use 2 subsets for training and 1 subset for testing

Assuming that we have the above training dataset (e.g., 14 samples, 3 features)

Step 1: Split training data to k folds

In this case let's assume k=3, so the training data will be split to 3 subset

feature1	feature2	feature3	prediction



Fold 1:
- training data: subset1 subset2
- Testing data: subset3

Fold 2:
- training data: subset1 subset3
- Testing data: subset2

Fold 3:
- training data: subset2 subset3
- Testing data: subset1

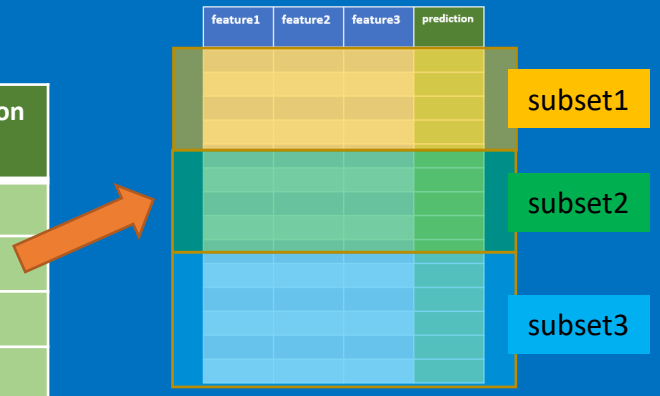
We can use 2 subsets for training and 1 subset for testing

Assuming that we have the above training dataset (e.g., 14 samples, 3 features)

Step 1: Split training data to k folds

In this case let's assume k=3, so the training data will be split to 3 subset

feature1	feature2	feature3	prediction



Fold 1:
- training data: subset1 subset2
- Testing data: subset3

Fold 2:
- training data: subset1 subset3
- Testing data: subset2

Fold 3:
- training data: subset2 subset3
- Testing data: subset1

We can use 2 subsets for training and 1 subset for testing

Step 2: train models and make prediction

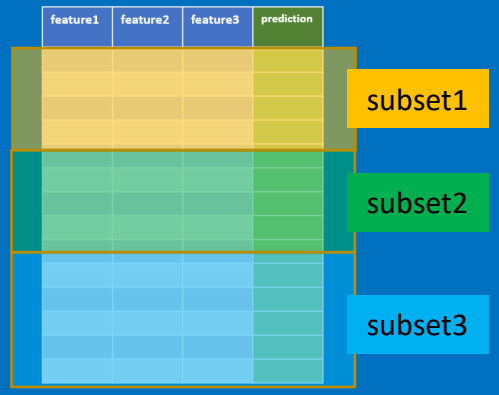
Assuming that we have the above training dataset (e.g., 14 samples, 3 features)

feature1	feature2	feature3	prediction

Assuming that we have the above training dataset (e.g., 14 samples, 3 features)

Step 1: Split training data to k folds

In this case let's assume k=3, so the training data will be split to 3 subset



Fold 1:
- training data: subset1 subset2
- Testing data: subset3

Fold 2:
- training data: subset1 subset3
- Testing data: subset2

Fold 3:
- training data: subset2 subset3
- Testing data: subset1

We can use 2 subsets for training and 1 subset for testing

Step 2: train models and make prediction

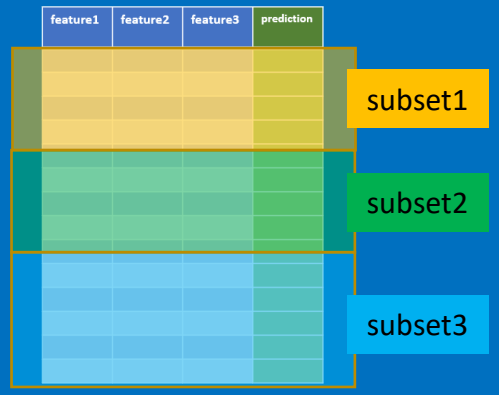
For each fold, we have training data and testing data

feature1	feature2	feature3	prediction

Assuming that we have the above training dataset (e.g., 14 samples, 3 features)

Step 1: Split training data to k folds

In this case let's assume k=3, so the training data will be split to 3 subset



Fold 1:
- training data: subset1 subset2
- Testing data: subset3

Fold 2:
- training data: subset1 subset3
- Testing data: subset2

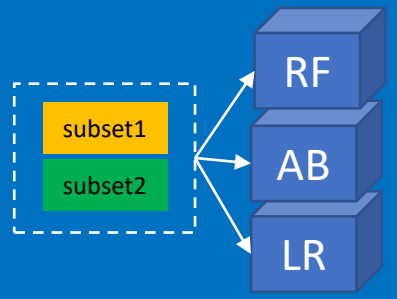
Fold 3:
- training data: subset2 subset3
- Testing data: subset1

We can use 2 subsets for training and 1 subset for testing

Step 2: train models and make prediction

For each fold, we have training data and testing data.

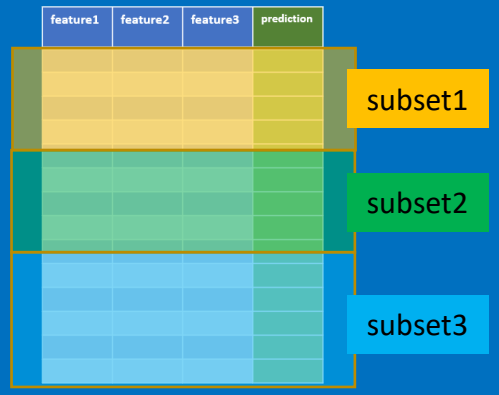
For example, for Fold1, we train 3 models: RF, AdaBoost and LR



Step 1: Split training data to k folds

In this case let's assume k=3, so the training data will be split to 3 subset

feature1	feature2	feature3	prediction



Fold 1:
- training data: subset1 subset2
- Testing data: subset3

Fold 2:
- training data: subset1 subset3
- Testing data: subset2

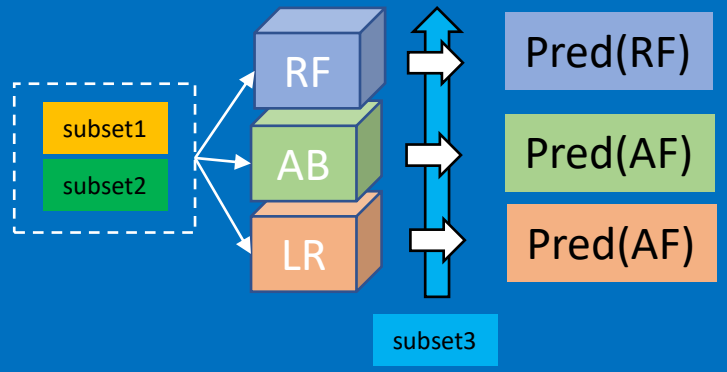
Fold 3:
- training data: subset2 subset3
- Testing data: subset1

We can use 2 subsets for training and 1 subset for testing

Step 2: train models and make prediction

For each fold, we have training data and testing data.

For example, for Fold1, we train 3 models: RF, AdaBoost and LR



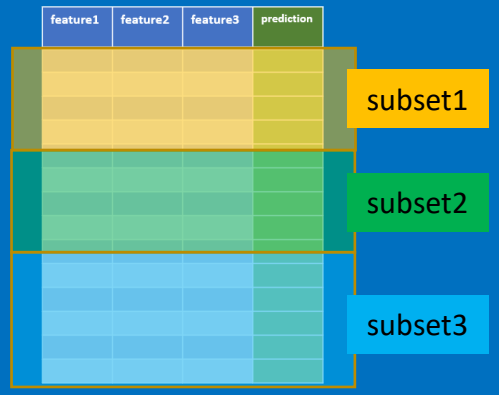
Assuming that we have the above training dataset (e.g., 14 samples, 3 features)

Then we apply testing data to the trained model and get the prediction

Step 1: Split training data to k folds

In this case let's assume k=3, so the training data will be split to 3 subset

feature1	feature2	feature3	prediction



Fold 1:
- training data: subset1 subset2
- Testing data: subset3

Fold 2:
- training data: subset1 subset3
- Testing data: subset2

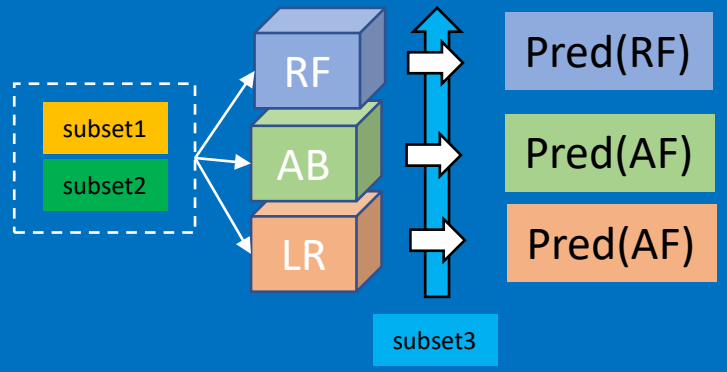
Fold 3:
- training
- Testing

We can use 2 subsets for training and

Step 2: train models and make prediction

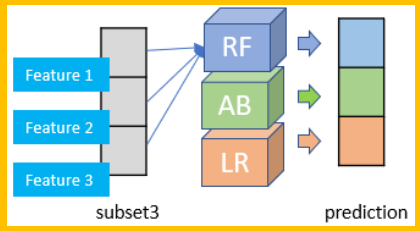
For each fold, we have training data a

For example, for Fold1, we train 3 mo



Then we apply testing data to the trained moc and get the prediction

Note that the predictions here are a matrix, for example, if subset3 only have one sample, then we have

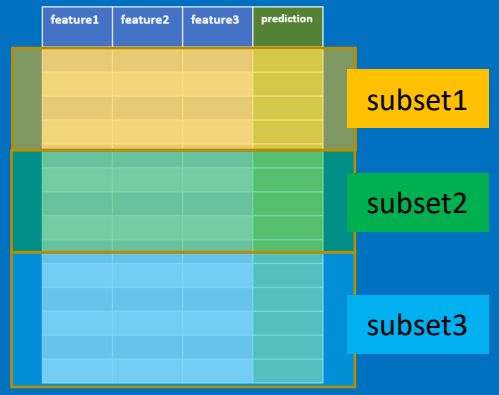


Assuming that we have the above training dataset (e.g., 14 samples, 3 features)

Step 1: Split training data to k folds

In this case let's assume k=3, so the training data will be split to 3 subset

feature1	feature2	feature3	prediction



Fold 1:
- training data: subset1 subset2
- Testing data: subset3

Fold 2:
- training data: subset1 subset3
- Testing data: subset2

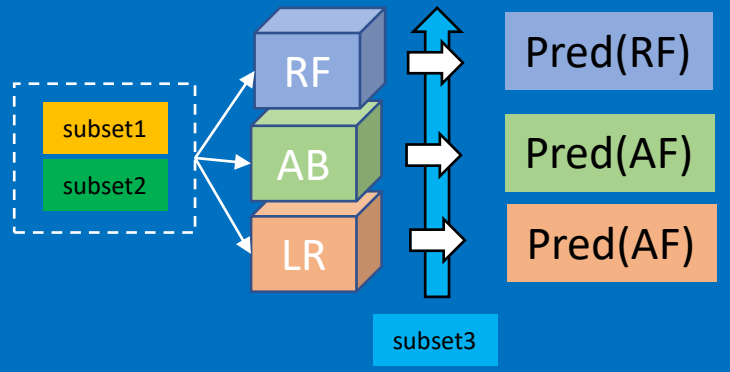
Fold 3:
- training
- Testing

We can use 2 subsets for training and

Step 2: train models and make prediction

For each fold, we have training data a

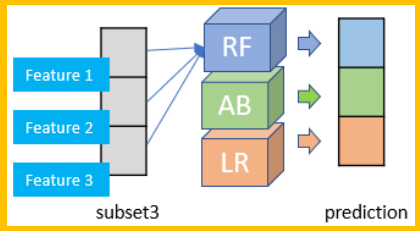
For example, for Fold1, we train 3 mo



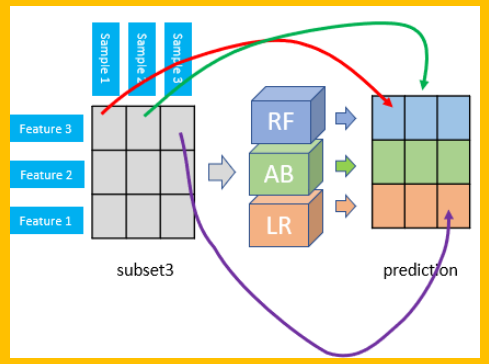
Assuming that we have the above training dataset (e.g., 14 samples, 3 features)

Then we apply testing data to the trained moc and get the prediction

Note that the predictions here are a matrix, for example, if subset3 only have one sample, then we have

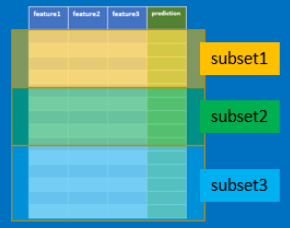


if subset3 have 3 sample, then we have the prediction as



Step 1: Split training data to k folds

In this case let's assume k=3, so the training data will be split to 3 subset



Fold 1:
- training data: subset1 subset2
- Testing data: subset3

Fold 2:
- training data: subset1 subset3
- Testing data: subset2

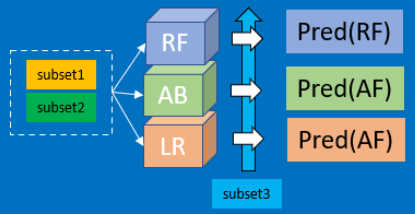
Fold 3:
- training
- Testing

We can use 2 subsets for training and

Step 2: train models and make prediction

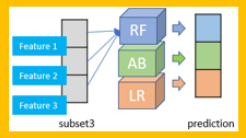
For each fold, we have training data a

For example, for Fold1, we train 3 mo

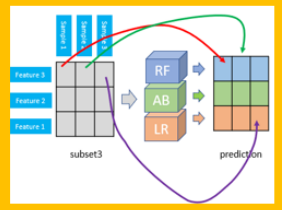


Then we apply testing data to the trained mod and get the prediction

Note that the predictions here are a matrix, for example, if subset3 only have one sample, then we have

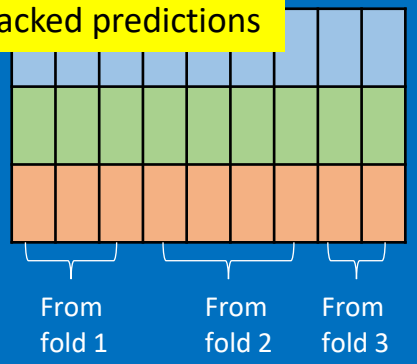


if subset3 have 3 sample, then we have the prediction as



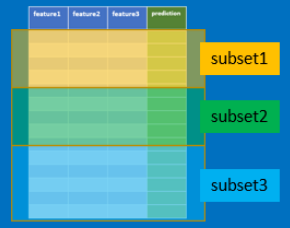
Step 3: we go through "step 2" for the rest fold2 and fold3, and we "stack" the predictions together

stacked predictions



Step 1: Split training data to k folds

In this case let's assume k=3, so the training data will be split to 3 subset



Fold 1:
- training data: subset1 subset2
- Testing data: subset3

Fold 2:
- training data: subset1 subset3
- Testing data: subset2

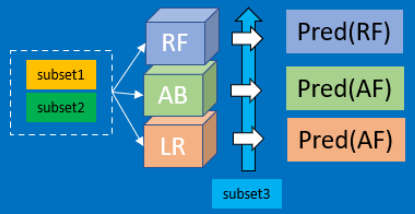
Fold 3:
- training
- Testing

We can use 2 subsets for training and

Step 2: train models and make prediction

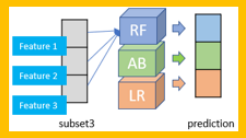
For each fold, we have training data a

For example, for Fold1, we train 3 mo

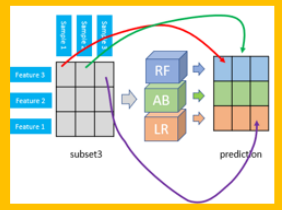


Then we apply testing data to the trained mod and get the prediction

Note that the predictions here are a matrix, for example, if subset3 only have one sample, then we have

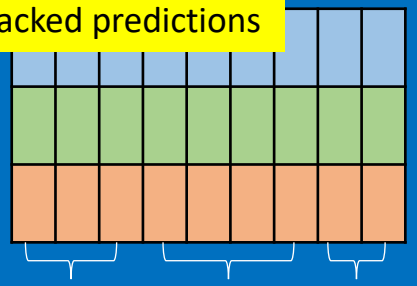


if subset3 have 3 sample, then we have the prediction as



Step 3: we go through "step 2" for the rest fold2 and fold3, and we "stack" the predictions together

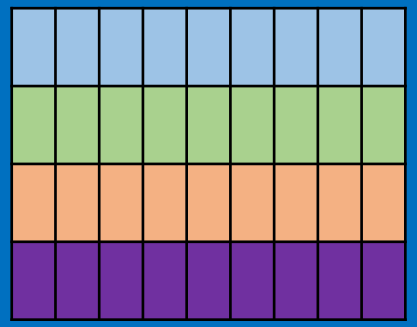
stacked predictions



From fold 1 | From fold 2 | From fold 3

Step 4: Then we attach the "truth" from the testing dataset to the predictions

truth ➡



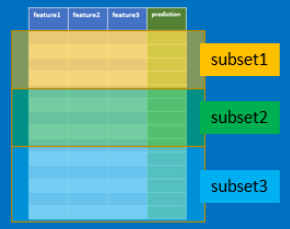
Prediction from RF

Prediction from AB

Prediction from LR

Step 1: Split training data to k folds

In this case let's assume k=3, so the training data will be split to 3 subset



Fold 1:
- training data: subset1 subset2
- Testing data: subset3

Fold 2:
- training data: subset1 subset3
- Testing data: subset2

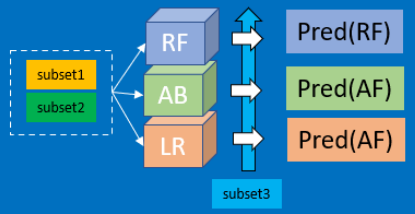
Fold 3:
- training
- Testing

We can use 2 subsets for training and

Step 2: train models and make prediction

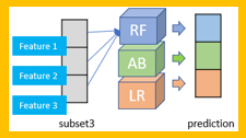
For each fold, we have training data a

For example, for Fold1, we train 3 mo

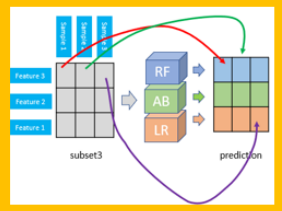


Then we apply testing data to the trained mod and get the prediction

Note that the predictions here are a matrix, for example, if subset3 only have one sample, then we have

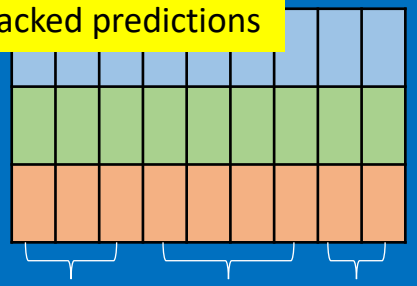


if subset3 have 3 sample, then we have the prediction as



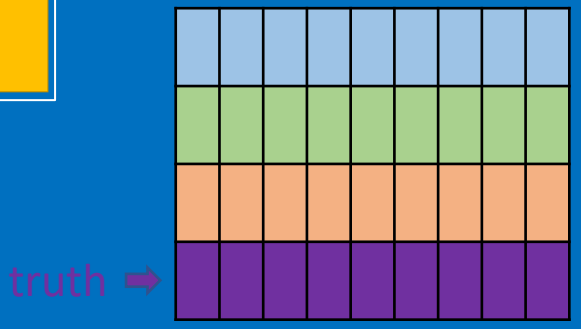
Step 3: we go through "step 2" for the rest fold2 and fold3, and we "stack" the predictions together

stacked predictions



From fold 1 | From fold 2 | From fold 3

Step 4: Then we attach the "truth" from the testing dataset to the predictions



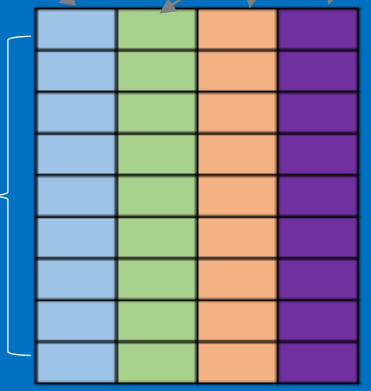
Prediction fr

Prediction fr

Prediction fr

Three features correspond to three models: RF, AB and LR

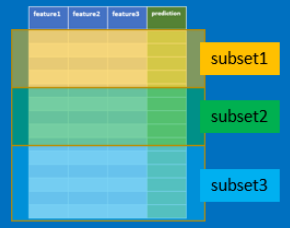
"truth"



Different samples

Step 1: Split training data to k folds

In this case let's assume k=3, so the training data will be split to 3 subset



Fold 1:
- training data: subset1 subset2
- Testing data: subset3

Fold 2:
- training data: subset1 subset3
- Testing data:

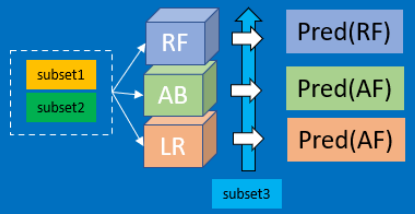
Fold 3:
- training
- Testing

We can use 2 subsets for training and

Step 2: train models and make prediction

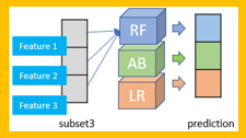
For each fold, we have training data a

For example, for Fold1, we train 3 mo

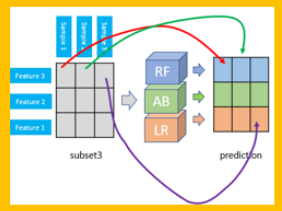


Then we apply testing data to the trained moc and get the prediction

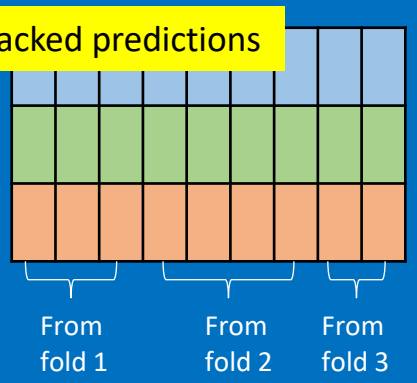
Note that the predictions here are a matrix, for example, if subset3 only have one sample, then we have



if subset3 have 3 sample, then we have the prediction as

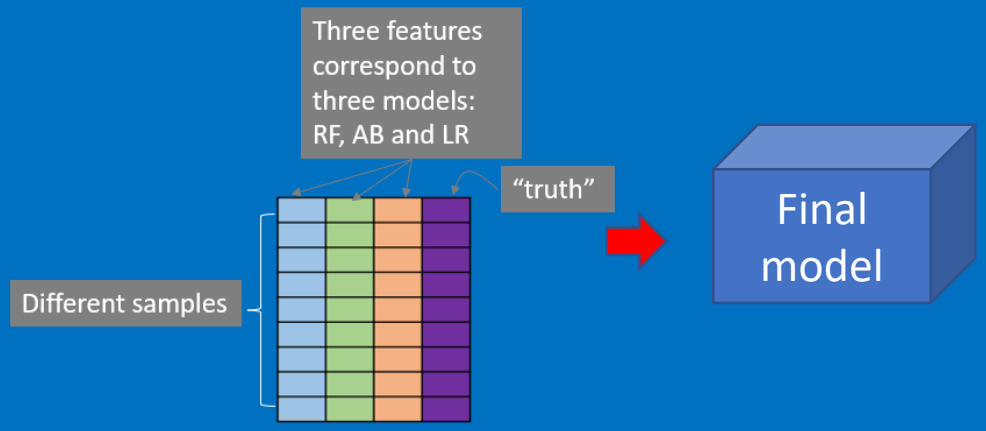


Step 3: we go through "step 2" for the rest fold2 and fold3, and we "stack" the predictions together



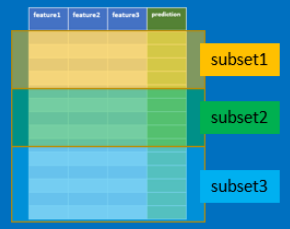
Step 4: Then we attach the "truth" from the testing dataset to the predictions

Step 5: We can use another model to "train" this new dataset



Step 1: Split training data to k folds

In this case let's assume k=3, so the training data will be split to 3 subset



Fold 1:
- training data: subset1 subset2
- Testing data: subset3

Fold 2:
- training data: subset1 subset3
- Testing data: subset2

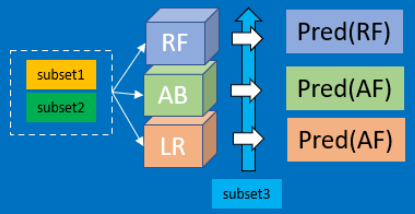
Fold 3:
- training
- Testing

We can use 2 subsets for training and

Step 2: train models and make prediction

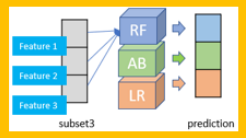
For each fold, we have training data a

For example, for Fold1, we train 3 mo

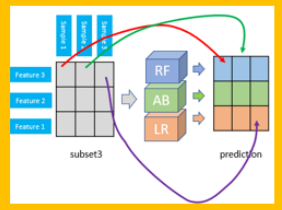


Then we apply testing data to the trained moc and get the prediction

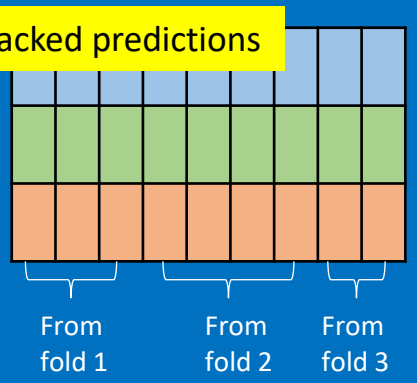
Note that the predictions here are a matrix, for example, if subset3 only have one sample, then we have



if subset3 have 3 sample, then we have the prediction as

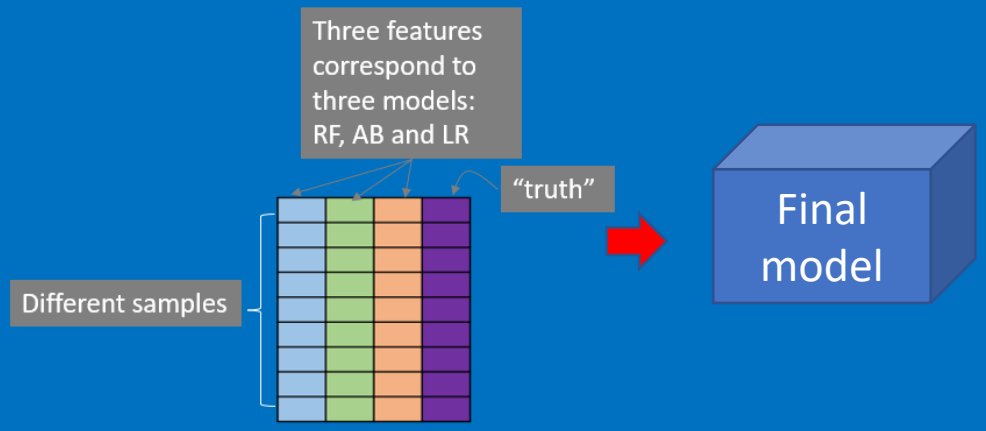


Step 3: we go through "step 2" for the rest fold2 and fold3, and we "stack" the predictions together



Step 4: Then we attach the "truth" from the testing dataset to the predictions

Step 5: We can use another model to "train" this new dataset



So if we have new testing data,

- first it will go through RF, AB and LR as usual
- Then we will have a prediction containing three individual predictions
- These 3 predictions will be treated as a completely new dataset, and being used in the "final" model to make the "final prediction"