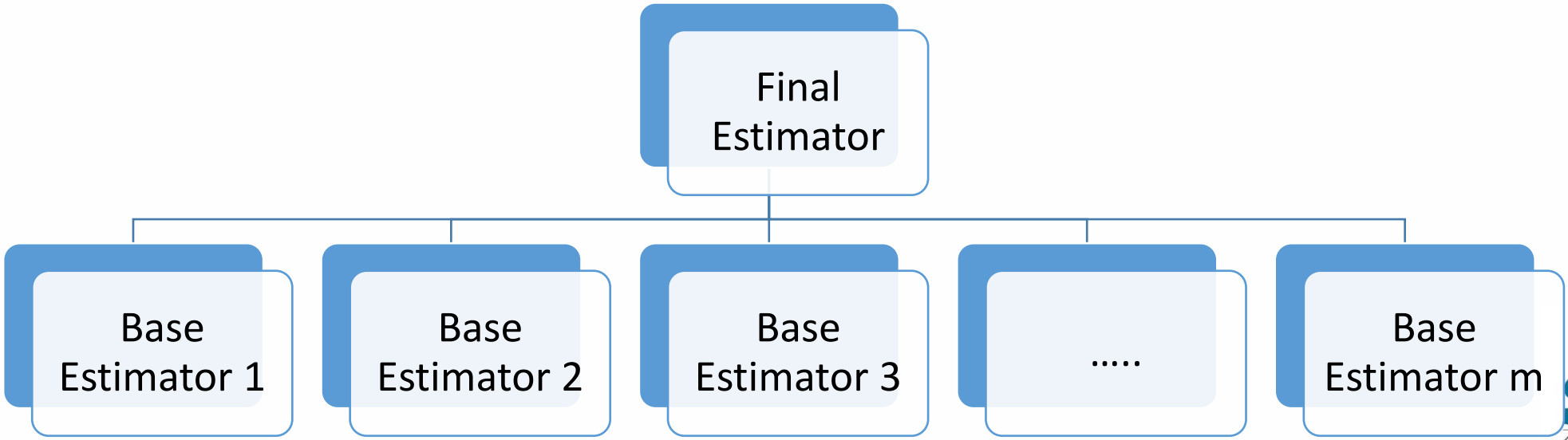
Sane's STATS Academy of Statistics

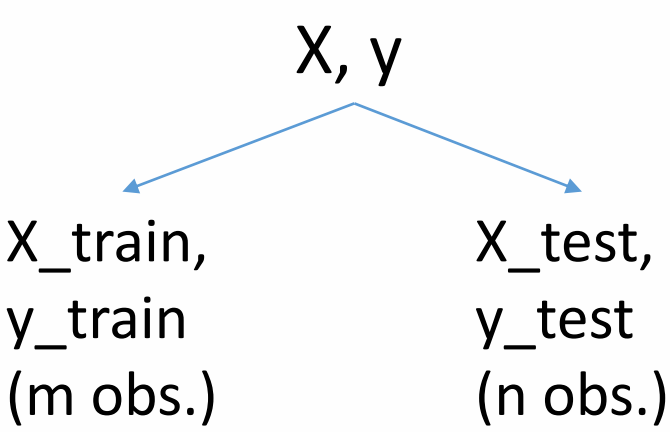
Stack Ensembling

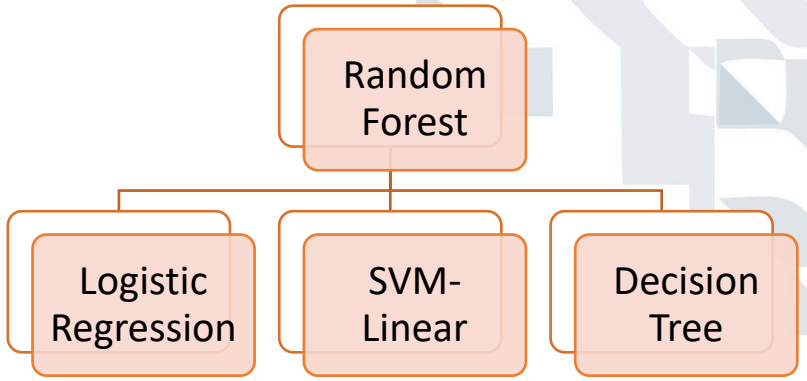
Stack Ensembling

- In stacking, a new model is trained from the predictions from various models
- Predicted columns act as features with response variable as the original one



Stack Ensembling Example : Train Set Operations





LogisticRegression().fit(X_train, y_train) & .predict(X_train) → y_pred_lr
SVC().fit(X_train, y_train) & .predict(X_train) → y_pred_svc
TreeClassifier().fit(X_train, y_train) & .predict(X_train) → y_pred_dtc

m obs

y_pred_lr	y_pred_svc	y_pred_dtc

X_trn_pred

m observations

y_train

m observations

RandomForestClassifier.fit(
X_trn_pred, y_train)

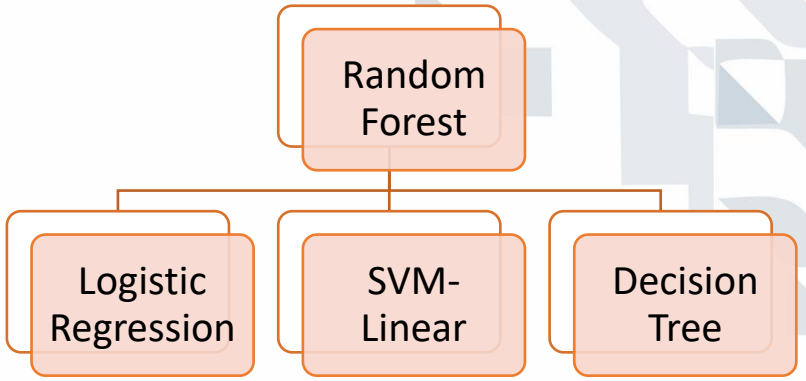
Sane's STATS Academy of Statistics

Stack Ensembling Example: Test Set Operations

X, y

X_train,
y_train
(m obs.)

X_test,
y_test
(n obs.)



LogisticRegression().predict(X_test) → y_pred_lr
SVC().predict(X_test) → y_pred_svc
DecisionTreeClassifier().predict(X_test) → y_pred_dtc

n observations

y_pred_lr	y_pred_svc	y_pred_dtc

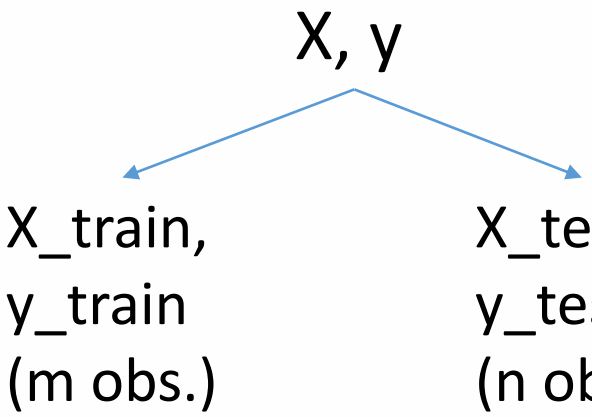
X_tst_pred

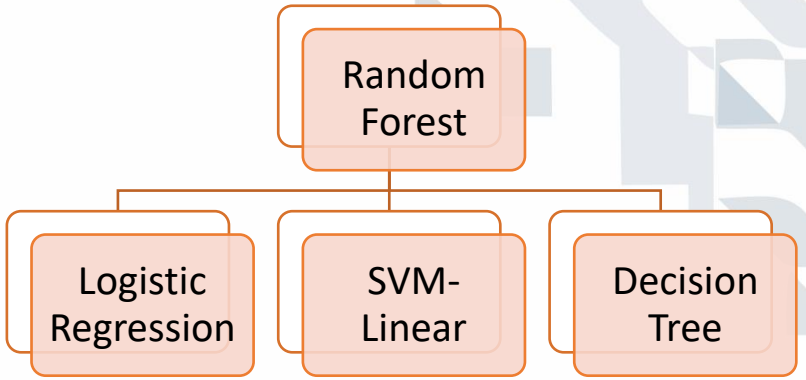
n observations

RandomForestClassifier.predict
(X_tst_pred)
→ Final Prediction

Sane's STATS Academy of Statistics

Stack Ensembling Example : Train Set Operations (Pass Through)





LogisticRegression().fit(X_train, y_train) & .predict(X_train) → y_pred_lr
SVC().fit(X_train, y_train) & .predict(X_train) → y_pred_svc
TreeClassifier().fit(X_train, y_train) & .predict(X_train) → y_pred_dtc

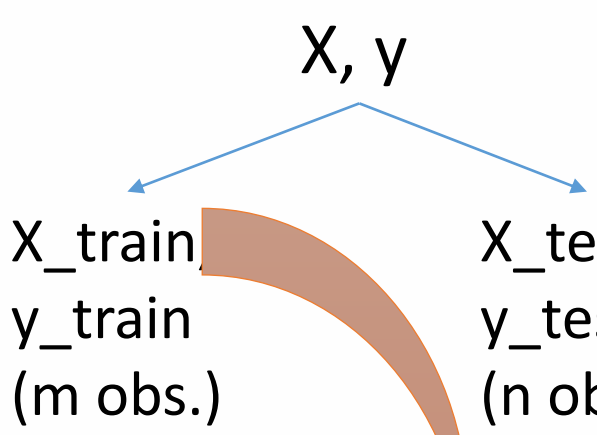
m obs

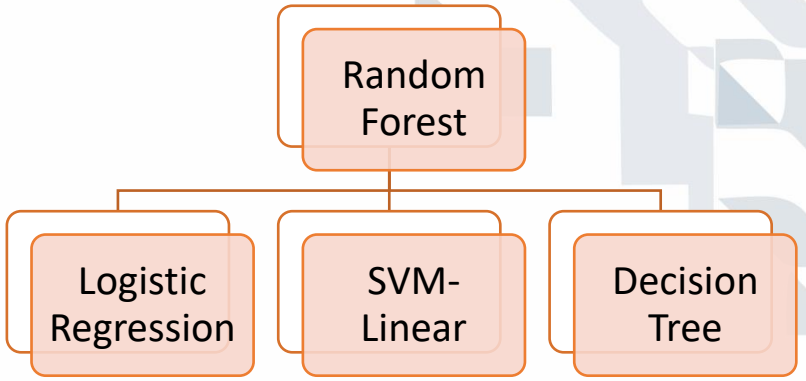
y_pred_lr	y_pred_svc	y_pred_dtc
	X_trn_pred	

m obs.

Sane's STATS Academy of Statistics

Stack Ensembling Example : Train Set Operations (Pass Through)





LogisticRegression().fit(X_train, y_train) & .predict(X_train) → y_pred_lr
SVC().fit(X_train, y_train) & .predict(X_train) → y_pred_svc
TreeClassifier().fit(X_train, y_train) & .predict(X_train) → y_pred_dtc

m obs

X_train	y_pred_lr	y_pred_svc	y_pred_dtc

X_trn_pred

m obs.

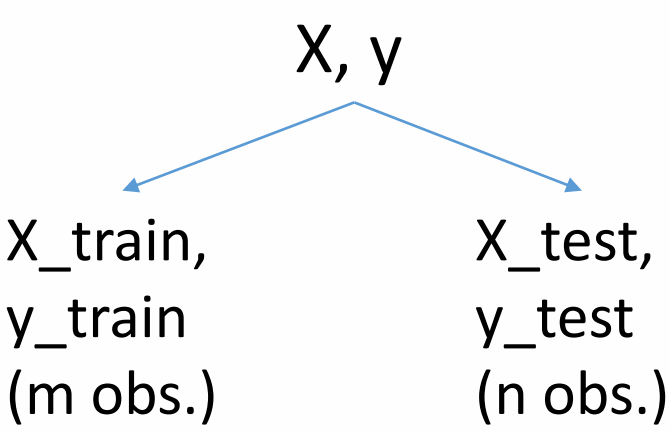
y_train

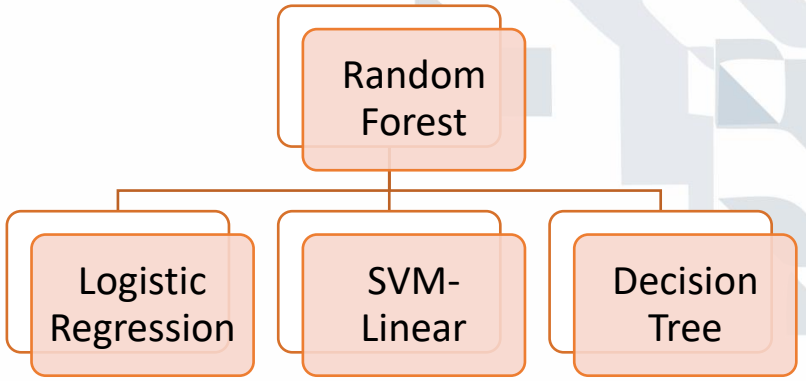
m obs.

RandomForestClassifier.fit(
X_trn_pred, y_train)

Sane's STATS Academy of Statistics

Stack Ensembling Example: Test Set Operations (Pass Through)





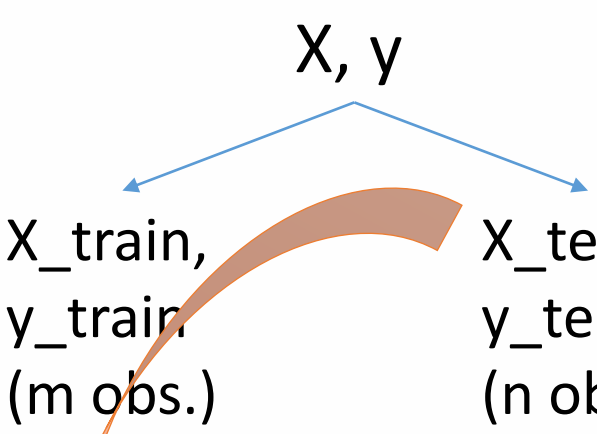
LogisticRegression().predict(X_test) → y_pred_lr
SVC().predict(X_test) → y_pred_svc
DecisionTreeClassifier().predict(X_test) → y_pred_dtc

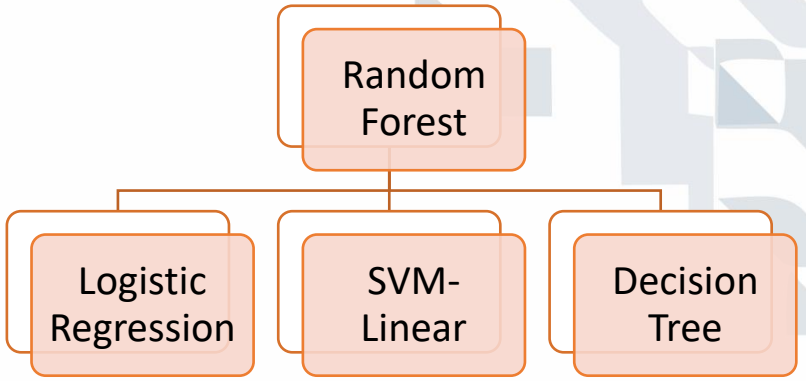
n observations

y_pred_lr	y_pred_svc	y_pred_dtc
	X_tst_pred	

n observations

Stack Ensembling Example: Test Set Operations (Pass Through)





LogisticRegression().predict(X_test) → y_pred_lr
SVC().predict(X_test) → y_pred_svc
DecisionTreeClassifier().predict(X_test) → y_pred_dtc

n observations

X_test	y_pred_lr	y_pred_svc	y_pred_dtc
X_tst_pred			

n Obs.

RandomForestClassifier.predict (X_tst_pred)
→ Final Prediction

Sane's STATS Academy of Statistics