DATA PROCESSING

STEP 1
DATA CLEANING

STEP 2
DATA HARMONIZATION
Data Sets - DS
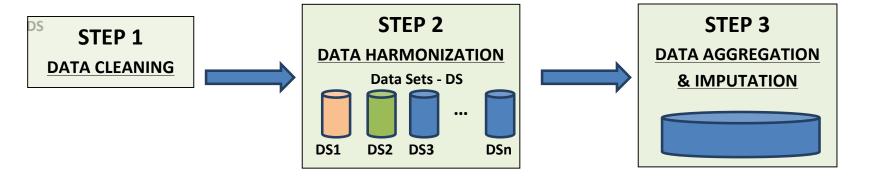DS1 DS2 DS3 ... DSn

STEP 3
DATA AGGREGATION & IMPUTATION

COMPRESSIVE BIG DATA ANALYTICS – CBDA ALGORITHM

STEP 4
SAMPLING FROM AGGREGATED DATASET

STEP 5
SUPER LEARNER – SL

STEP 6
Knockoff – KO

STEP 7
K SETS OF RESULTS

SL SL SL SL SL .... SL

1 2 3 4 5 .... K

STEP 8
MODEL MINING or SYNTHESIZER