

Steps usual in any predictive analysis:

1. Data description
 - a. Description of type of data and variables
 - b. Sample number
 - c. Variable number
 - d. If it is classification problem: Table of number of samples by class
 - e. If it possible, do a short description of data:
 - i. Multivariate analysis: Principal Component Analysis (PCA), clustering. ...
 - f. What is the type of predictive problem?
 - i. Classification / Regression
2. Preprocess?
3. Partition of data?
 - a. Train/test
 - b. Train: k-fold crossvalidation (repeated), bootstrap, ...
4. Train model (tuning parameters?)
5. Predict test data
 - a. Evaluation metrics.
 - i. Classification problem
 - ii. Regression problem
6. Comparison models
7. Choose best model