

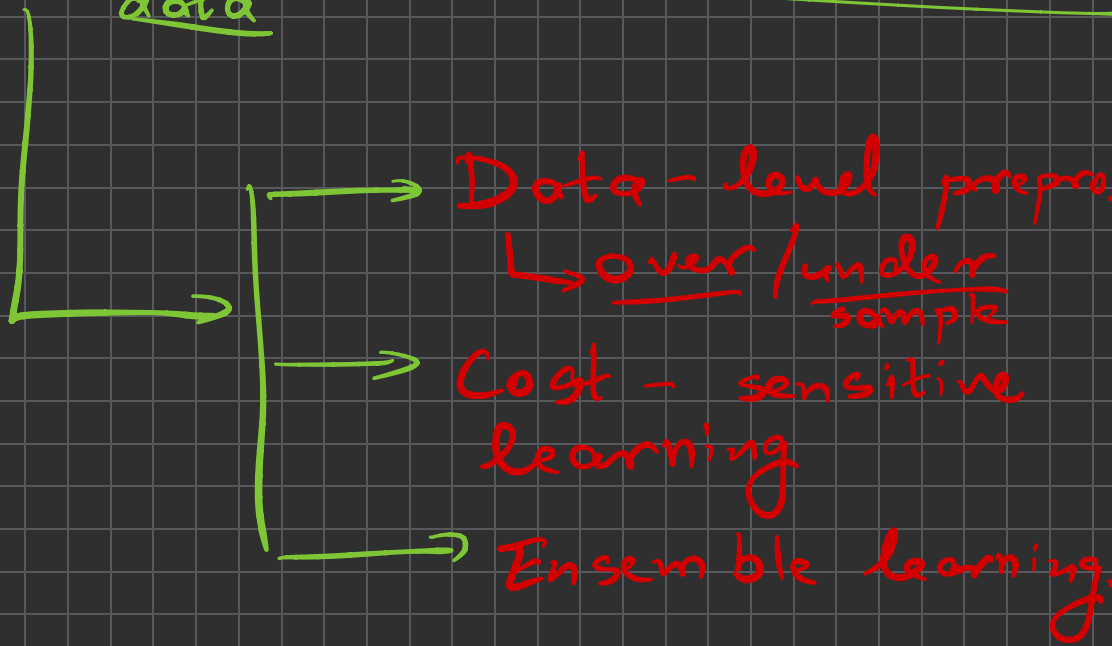
Possible problem statements

→ reduce no. of classes in target variable into smaller no. of classes to draw a more comprehensive classification

→ a direct classification and another classification to show difference of PCA, class imbalance etc.

→ Use ensemble methods
to create an
even better
classification
model.

Learning from imbalanced data



→ over sampling using random sampling

→ use imblearn.
oversampling

RandomOverSampler

→ Report flow

① Introduction

② Development of Questions

③ Analysis Strategy and Methodology.

→ Strat.

→ pre.

→ vis.

→ SMOTE

→ Experiment

→ PCA

④ models shortlisted

⑤ Result analysis → (Not long and to point)

⑥ Analysis Code snippets.