STRUCTURE DIAGRAM

prepare train and test
datasets from each
collection for each class

spam

ham

info

nets

test

SpamPattern
(particular class rules)

InfoPattern
(particular class rules)

NetsPattern
(particular class rules)

Base Pattern class
(common rules)
pattern_wrapper.py

HamPattern
(particular class rules)

msg_wrapper.py

vectorizer.py

DMARG Checker

Destinator Checker

Originator Checker

Subject Checker

List Checker

Emarket Checker

Content Checker

URL Checker

Attach Checker

MIME Checker

Base Checker

checker.py

sklearn.normalize.normilize()

sklearn.select
SelectKBest (ANOVA f_score)

sklearn.ensemble
RandomForestClassifier

sklearn.ensemble.
ExtraTreesClassifier

cls_wrapper.py

estimated SPAM P %

estimated HAM P %

estimated INFO P %

estimated NETS P %

merge results for
each email from
test subdir in report

estimated SPAM P %

estimated HAM P %

estimated INFO P %

estimated NETS P %

Report with
emails
statuses