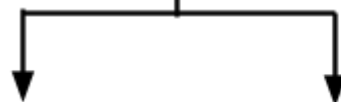


Step 1: Labelled data,
57502 entries, 38972
allowed, 18530 denied.

Step 4: Non inclusion of
the URL core domain
feature, leaving the URL
TLD for classifying.

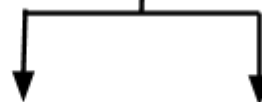
Unbalanced

Balanced (Undersampling &
Oversampling)



Random

Consecutive



Random

Consecutive