Detecting Malicious Twitter Users Using Mixed Features

1.0 Data Preprocessing

- Type conversions, tweet-document corpus creation

2.0 Feature Engineering

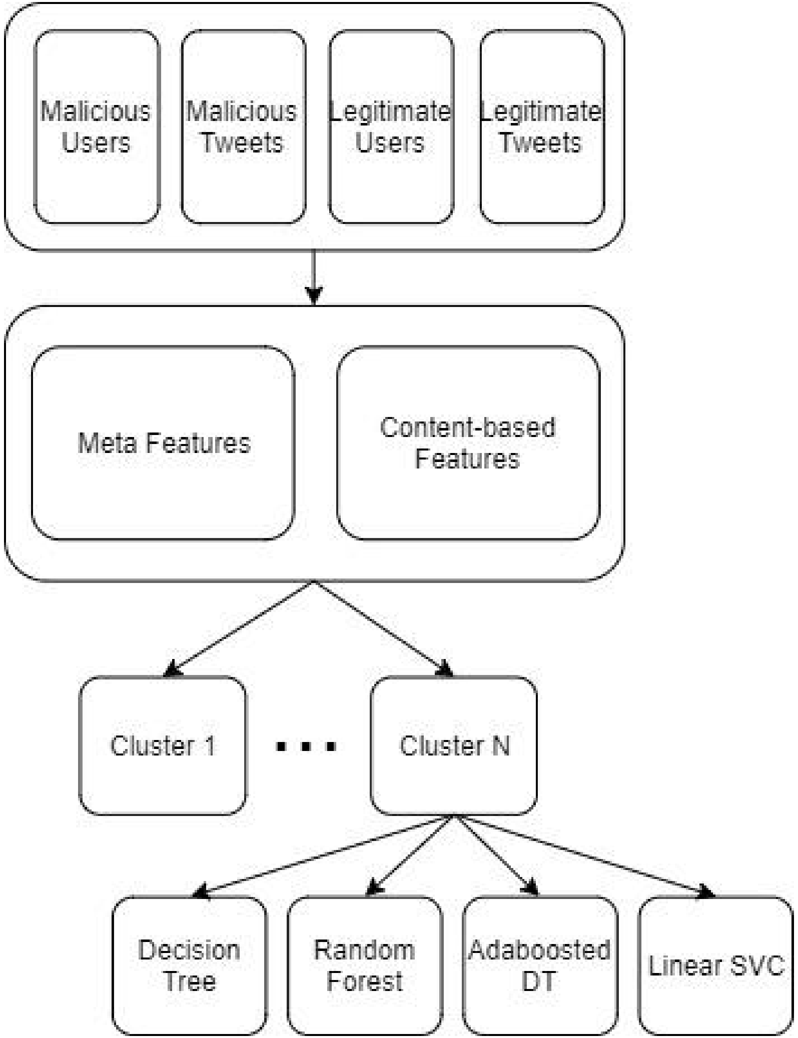
- Meta features (num. Followers, following)
- Content-based, NLP features (GOSS, LOSS, Document-topic distribution entropy)
- Latent Dirichlet Allocation distribution

3.0 Preliminary Clustering

- Kmeans clustering => data frame segmentation
- Focused, balanced, unfocused users

4.0 Classification

- Decision Tree, Random Forest, Adaboosted DT, Linear SVC
- Applied to all cluster segments
- K-fold CV accuracy, precision, recall and F1 scores


Malicious Users
Malicious Tweets
Legitimate Users
Legitimate Tweets
Meta Features
Content-based Features
Cluster 1
. . .
Cluster N
Decision Tree
Random Forest
Adaboosted DT
Linear SVC