ReDAS Lab

Phishing Detection with Deception Features

Victor Zeng

Advisor: Rakesh M. Verma

COMPUTER SCIENCE

Problem

- Phishing is the act of sending fake emails to trick a user into doing something.
  - Beachhead for 95% of attacks on enterprise networks
  - Average cost: $1.6 Million
- Cannot depend on user to identify phishing emails
- Creating labeled training data is expensive

Source: Eitan Katz. Phishing statistics: What every business needs to know, May 2019

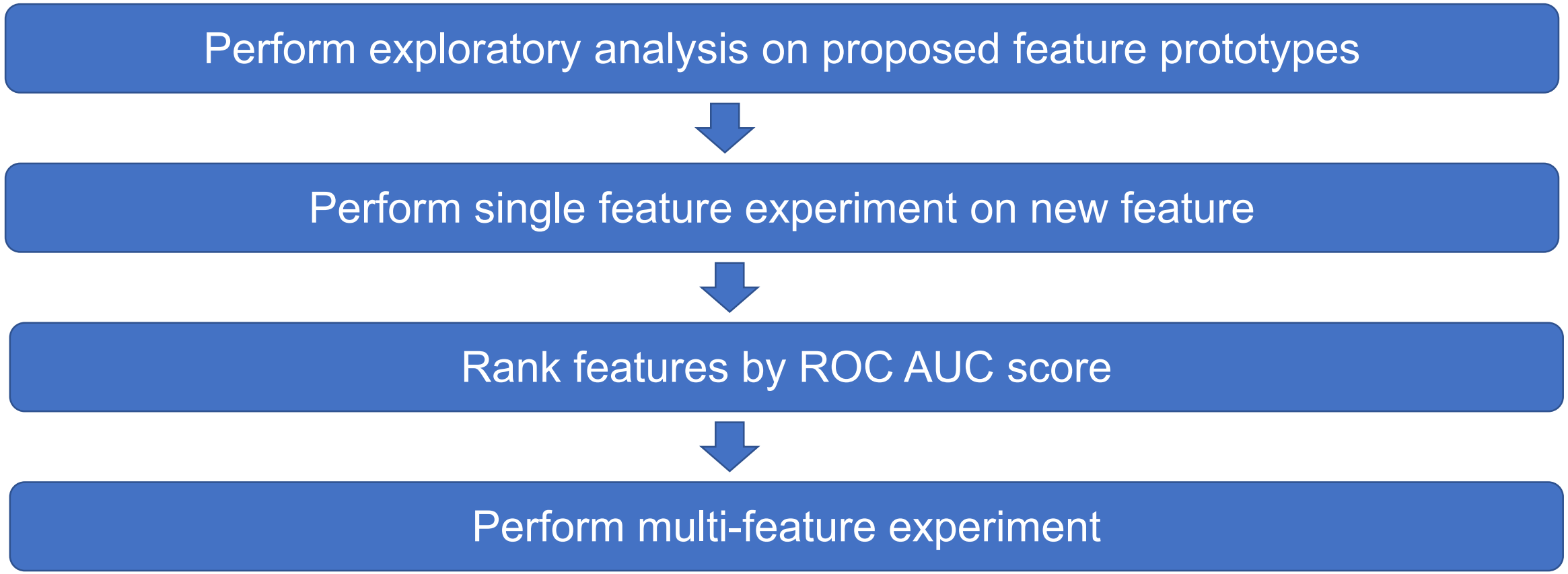
Project Overview

- Goal: Improve automated phishing detection
- Objective: Identify new features which can be used for phishing detection
- Expected Impact: Improve performance of email filters

Accomplishments

- Identified 6 new features which can be used in the detection of phishing emails.

Method Overview



Deliverables

- Code + Documentation for new features
- Poster
- Final Presentation
- Write-up

Methods: Exploratory Analysis

1. Prototype feature extraction code
2. Run feature extraction on IWSPA Dataset
3. Generate plots of feature values
4. Calculate kurtosis of feature results
5. Perform a two-sample t-test on feature values

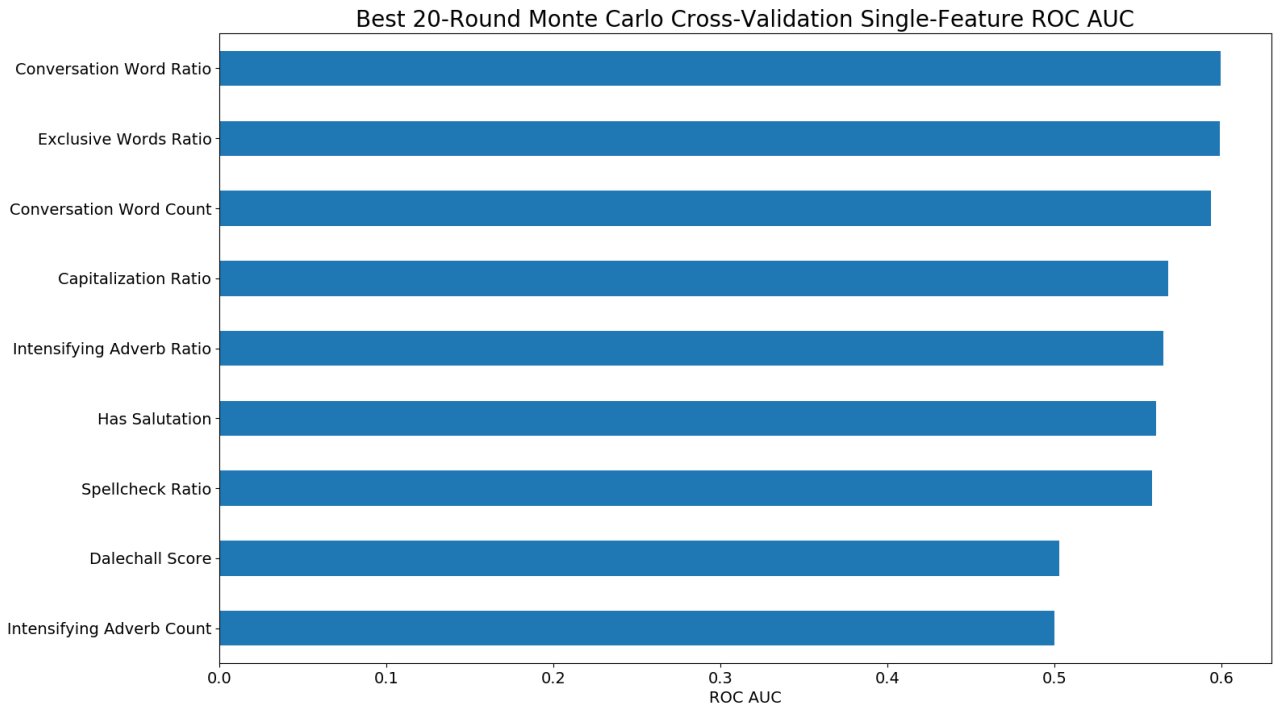
Methods: Single Feature Experiments

- Implement Feature in PhishBench
- Perform 20-round Monte-Carlo Cross Validation
  - Training set of 1000 Legit emails and 100 phish emails
- Dataset: IWSPA
- Classifiers Used: Bagging, Boosting, Decision Tree, 3NN, Logistic Regression, Multinomial Naïve Bayes, Naïve Bayes, Random Forest, SVM

Results: Exploratory Analysis

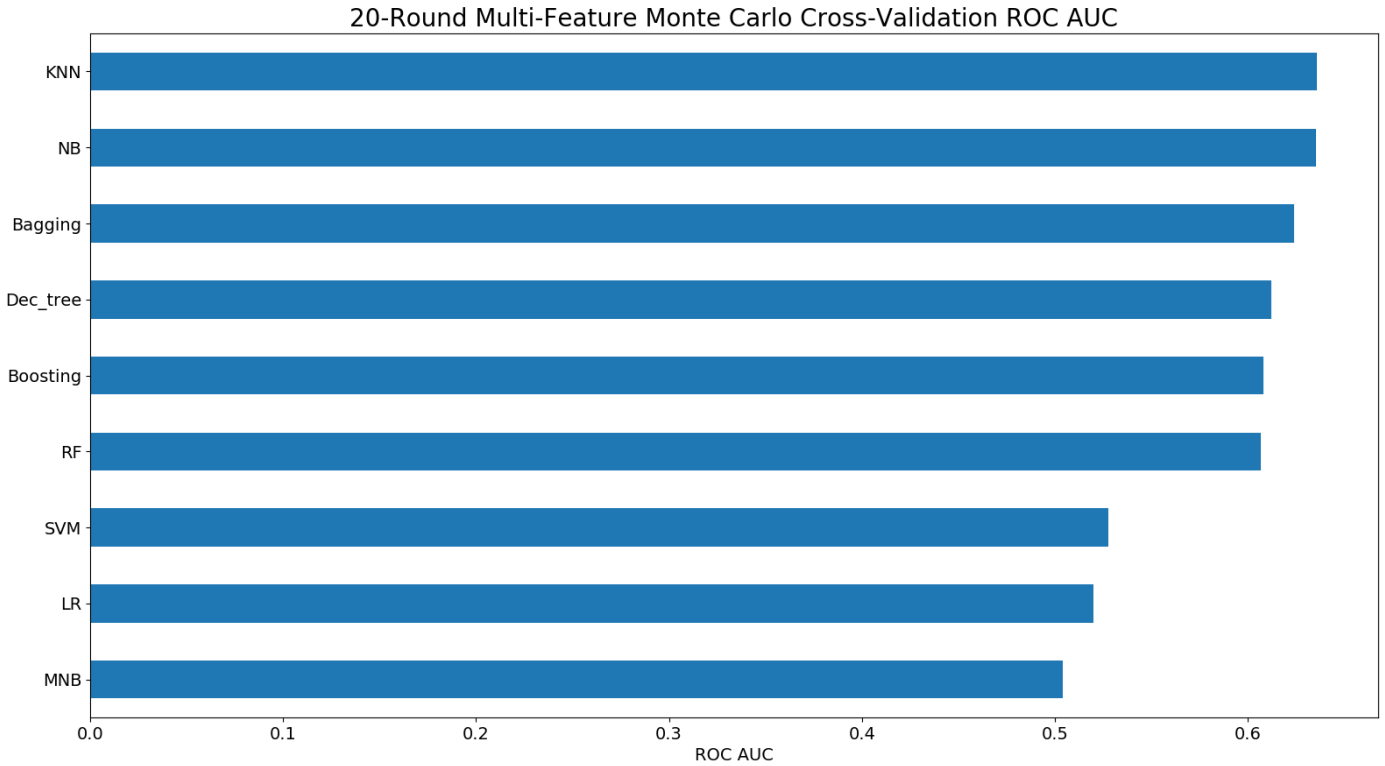
- Tested 8 features, each with several variants
  - 7 Features produced p-values under 0.0001 threshold
- Many features had very small p-values in the t-test

Results: Single Feature Experiments



- Most features produced ROC AUC scores between 0.56 and 0.6
- Counting features performed better when normalized
- Best performing classifier was normally random forest.

Results: Multi-Feature Experiments



- Most classifiers achieved ROC AUC over 0.6
- Multinomial Naïve Bayes, Logistic Regression, and SVM performed poorly

Conclusions

- Several deception features are useful for detecting phishing emails.
- The deception features are non-redundant.

Limitations

- The features identified in this research by themselves are not enough to accurately perform phishing detection.
  - Highest ROC AUC of multi-feature test is 0.635636
  - Will need to be augmented with additional features

Future Work

- Prototype and test additional features to use for phishing detection
- Test features in conjunction with traditional phishing detection features.
- Acquire a second dataset and perform cross-dataset validation of features
- Create general method for identifying key words in email corpus to serve as features.

Acknowlegements

The REU project is sponsored by NSF under award NSF-1659755. Phishing research in the ReDAS lab is made possible by an NSF grant to the University of Houston Computer Science Department (NSF CNS 1319212). Special thanks to the following UH offices for providing financial support to the project: Department of Computer Science; College of Natural Sciences and Mathematics; Dean of Graduate and Professional Studies; VP for Research; and the Provost's Office. The views and conclusions contained in this presentation are those of the author and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the sponsors.