A Natural Language-Based Phishing Detection and Risk Prevention System

Elena Shen, Christine Zhou, Ya-Fei Lin, Matthew A. Lanham
Purdue University Krannert School of Management
shen462@purdue.edu; zhou977@purdue.edu; lin1170@purdue.edu; lanhamm@purdue.edu

ABSTRACT

We researched and developed a natural language-based phishing detection and risk prevention system that allows one to accurately predict if an email is phishing email. The motivation for this work is the rise in cyber-attacks via malware or ransomware, particularly against company's environments. Our approach provides an extension app that alerts the user when a particular email shows signs of phishing risks. Users have the ability to report the email. With the data gathered, the app is able to predict similar phishing emails and warn users ahead of the time. The more data gathered, the higher the accuracy of the prediction rate. We achieved 97.31% AUC. This work was presented in the recent Crossroads Classic Analytics challenge among Butler University, Indiana University, Purdue University, and the University of Notre Dame where we won first place in the undergraduate division.

INTRODUCTION

In 2020, 74% of US companies experienced security breaches via phishing attacks and cost companies up to $14.8 annually (Rosenthal, 2020). Phishing is when an unauthorized third party sends viruses in forms of fraudulent links, ads, emails, and text messages to steal personal information and data. The motivation of this study is to prevent and protect against these threats as it causes huge repercussions such as: intellectual property theft, operational disruption, and financial losses. Of the 74% US companies, 60% of the organizations lost data and 18% experienced financial losses (Vaas, 2022).

Our research tackles the detection of phishing emails in employee inbox and reducing data breaches. We use three different predictive models to differentiate phishing emails from regular functional emails. The research team hypothesizes the Naïve Bayes Method will provide the most accurate phishing detection results because it can efficiently categorize and make predictions based on the given data. It will classify phishing with non-phishing emails.

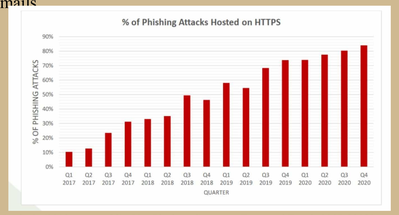


Fig 1. Cook, S. (2018, August 28). *50+ Phishing Statistics, Facts and Trends 2017-2018 | Comparitech*. Comparitech. https://www.comparitech.com/blog/vpn-privacy/phishing-statistics-facts/

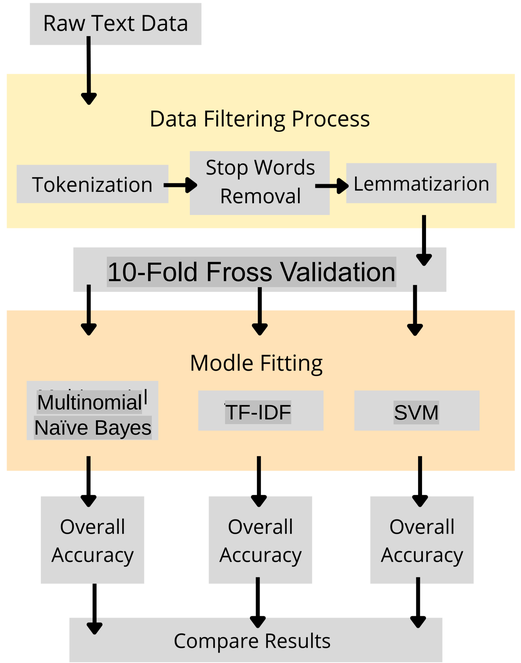
RESEARCH OBJECTIVES

- RQ1: How effective are predictive models in accurately detecting phishing emails?
- RQ2: How can we use natual language on phishing email to reduce business risk?
- RQ3: Can we use our analytics skills to develop a tool to minimize phishing email click-rate?

LITERATURE REVIEW

- (2020, Zamir) Phishing Web Site Detection Using Diverse Machine Learning Algorithms
  - Contribution: Stacking different predictive models
- (2019, Fang et al.) Phishing Email Detection Using Improved RCNN Model With Multilevel Vectors and Attention Mechanism
  - Contribution: Focused on one model only
- (2018, NB et al.) A Machine Learning Approach Towards Phishing Email Detection CEN-Security
  - Contribution: Genral overview of how machine learning is used for phishing email detection

METHODOLOGY



Click to add text

STATISTICAL RESULTS

After running 10 k-folds cross validation, we have found that even though the multinomial Naïve Bayes model produces the highest average accuracy of **98.4%**, it has an outlier prediction with accuracy of 96.3%. TF-ID's 25-75 percent predictions are higher compared to the other 2. SVM's prediction averages the lowest, but it has the smallest standard deviation, meaning it is more precision.

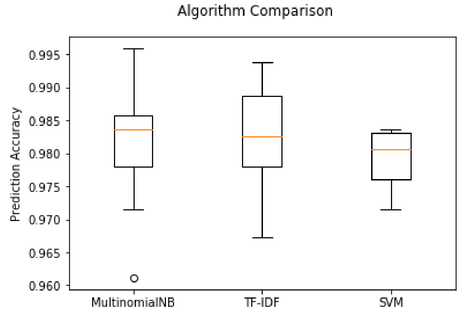


Fig 2. Algorithm comparison with 10 k-folds cross validation

EXPECTED BUSINESS IMPACT

Based on the statistical results above, we used the multinomial Naïve Bays method to evaluate business impact since it yields higher overall predicting accuracy. On average, phishing emails cost businesses 15 million dollars per year. With our phishing detection app that uses the multinomial Naïve Bays machine learning method, business **could potentially save $14,441,000**.

Figure 3 demonstrates the overall business savings if a company decides to use our phishing detection product. Our algorithm predicted 98% actual phishing emails as true and only 1.8% are false positives. When combined with the information above, we predict that our solution will save businesses 14.72 million dollars, minus the false positive predictions we get 14.44 million saving.

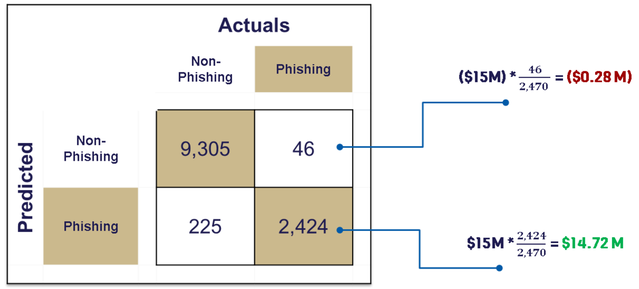
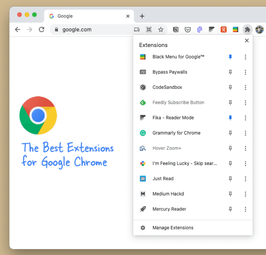
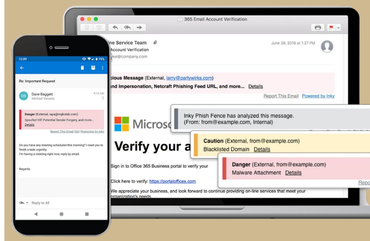


Fig 3. Total Net Savings ($) based on predicted phishing emails

DEMONSTRATION

- The extension can be downloaded to any email platform with the purpose of detecting phishing emails.
- Notification alerts are color coded, indicating the likelihood of a phish-risk. Users have the option to report & block email.
- Machine learning: constantly collecting data to improve predictive accuracy.
- By using our mockup app:
  - Decrease in phishing click rates leads to decrease in phishing attacks.
  - Businesses will be protected against financial and data loss.





- ***90%***=Automatic block
- ***70%***=Red Flag
- ***50%***=Yellow Flag
- ***30%***=Gray Flag

CONCLUSIONS

- Through our research, we discovered that **Naïve Bayes** model yields the highest accuracy overall with an **accuracy of 98.4%**.
- By using the multinomial Naïve Bayes model for the phishing detection, there will be an increase in data breach prevention among business companies, **saving $14.44 million per company**.
- **We developed a prototype extension app** with the purpose of decreasing the negative impact of phishing attacks. The primary function is to detect emails that contains phish-risks and notify the user to block and report. The prototype will contribute to bringing more knowledge and understanding to employees regarding phishing.
- In the future we will consider publishing this work in **The Journal of Purdue Undergraduate Research** after we investigate more models used in previous studies to fill the research gaps currently in this area.

ACKNOWLEDGEMENTS

We would like to thank business leaders at IN3, ECS Tech, and Midcontinent Independent System Operator (MISO) for providing us this opportunity. We would also like to thank Professor Matthew Lanham for guiding us throughout this project.

PURDUE UNIVERSITY
Krannert School of Management

CROSSROADS CLASSIC