# Replication and Extension of Advanced Phishing Detection:

# Integrating Machine Learning with Email Services for Malicious URL Identification

Tyler Beasley, Keshawn Blakely, Matthew Tieman

# Problem Statement

The rate and complexity of cyber-crimes are constantly increasing, with phishing attacks being a primary concern. Meaning more robust methods are necessary for effectively identifying these attacks.

**Challenge**: Develop a phishing detection system based on tokenization and multimodal features using RNN-based models for accurate and efficient classification of URLs being benign or malicious.

# Technical Challenges

While the motivation and outcome of this project is relatively straightforward, difficulty arises during various steps of the technical integrations. Among the many challenges, the primary ones include:

- Properly training the RNN model
- Yielding high-accuracy results
- Appropriately integrating the model with a(n) email service(s)

Overall, these challenges become difficult because these attacks are continuously evolving. So, while the training may yield great results, using the application with real-world URLs, results may vary drastically.

# Related Work

Our project is intended to be a replication and extension of the existing work, "Multimodal Phishing URL Detection Using LSTM, Bidirectional LSTM, and GRU models" where we utilize the proposed models and extend on their work by integrating it with email services practical usage of these models as the forementioned work focuses on.

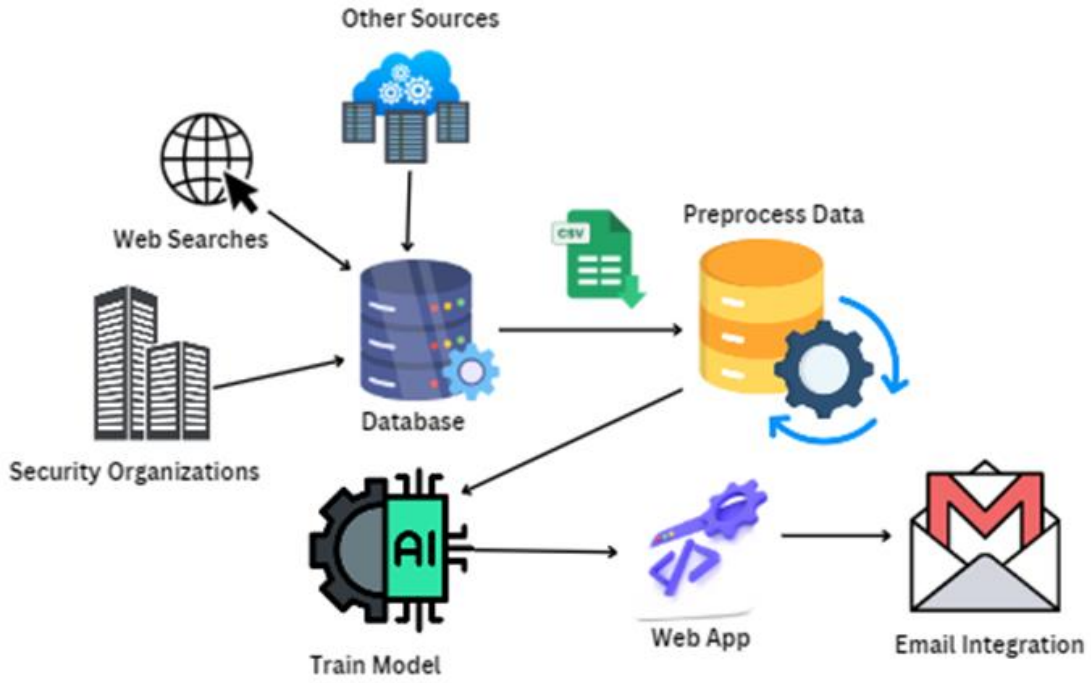
Our project aims the address the practical usage of these models as the aforementioned work focuses on the architecture and evaluation of these models.

# Approach

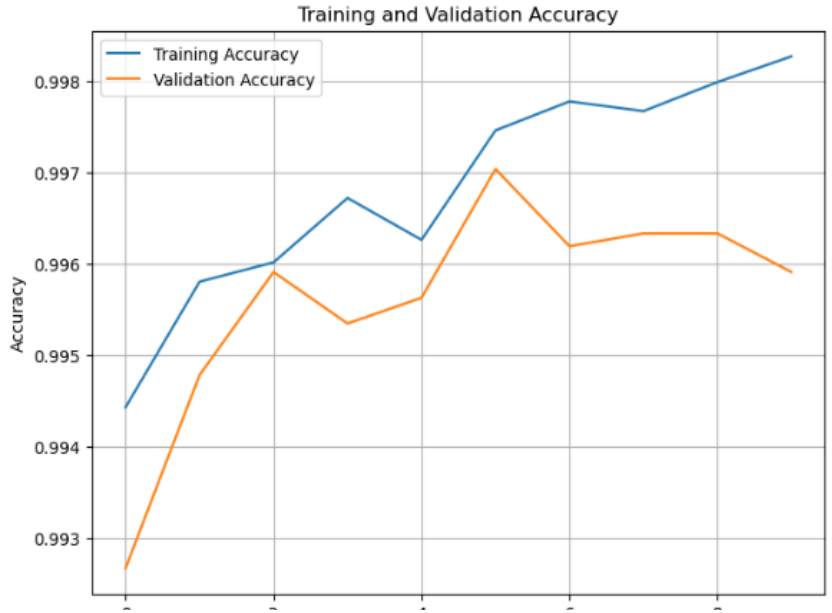
Our approach follows a relatively simple process:

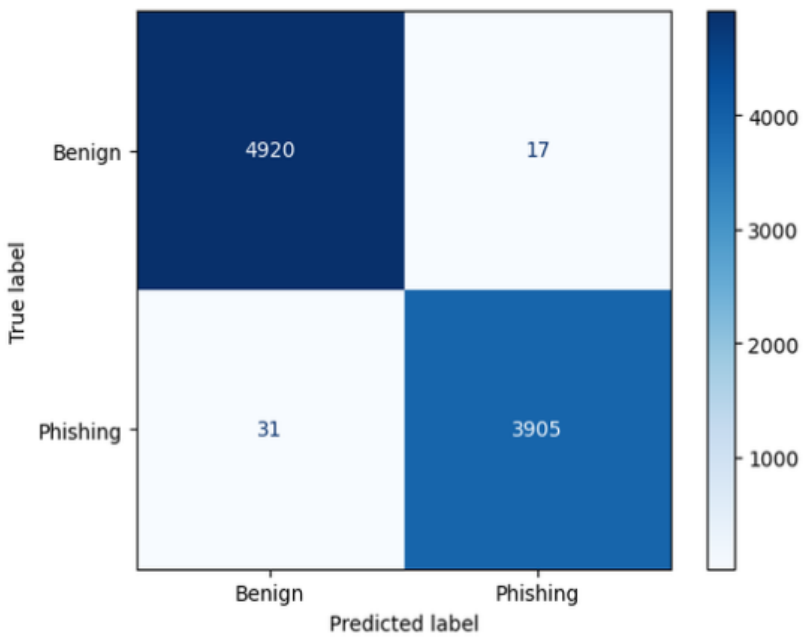
- Obtain the dataset
- Train the RNN models
- Replicate performance results
- Integrate the model with an email service utilizing Google API

We used a dataset provided by PhishTank where they collect an abundance of both malicious and benign URLs. To measure the performance, we used accuracy tests, training and validation graphs, and confusion matrices.

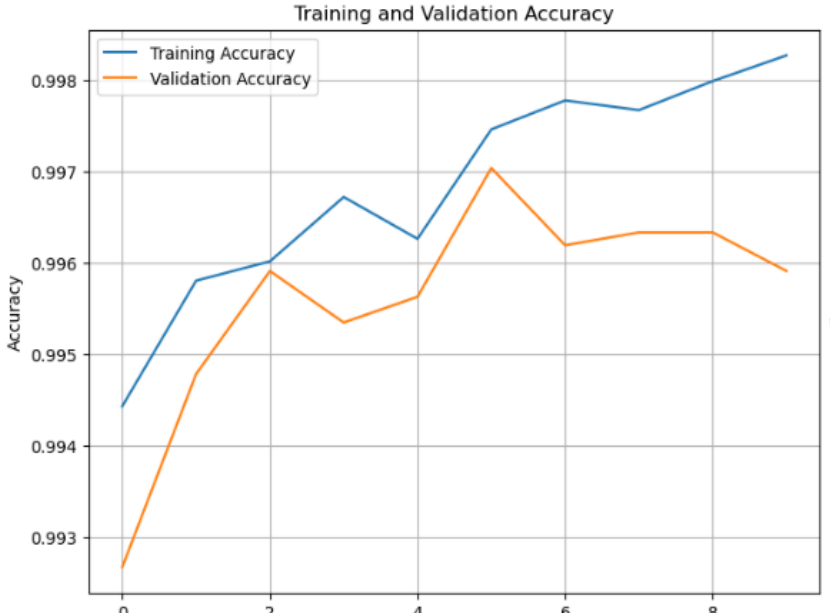


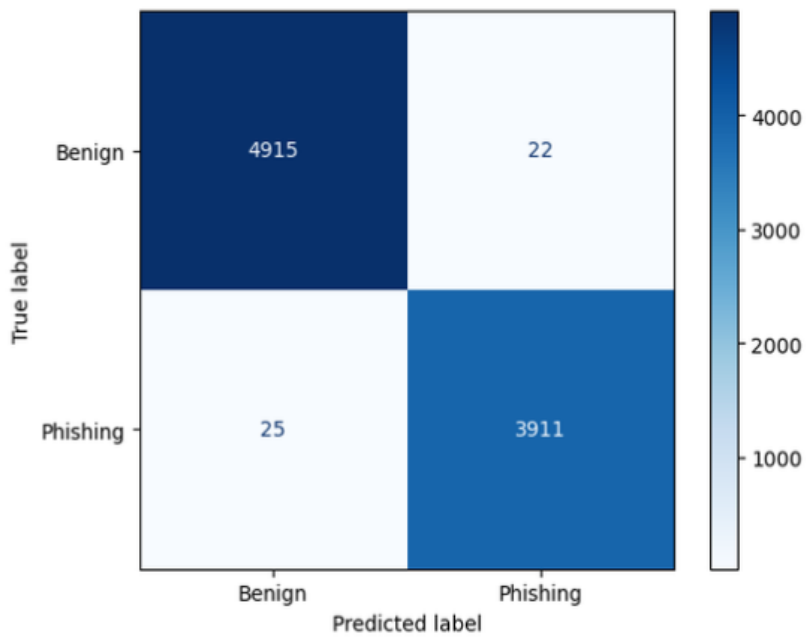
LSTM (99.45%)



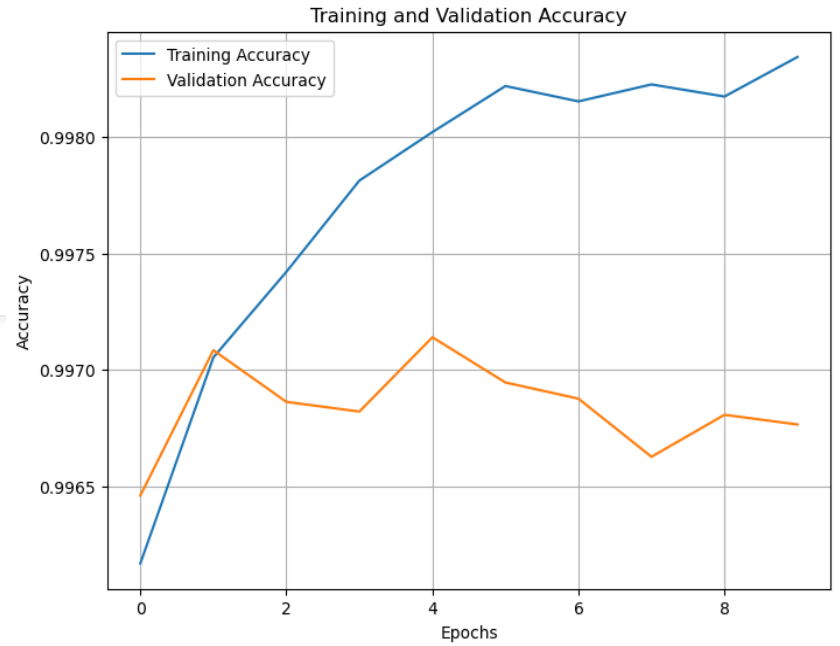



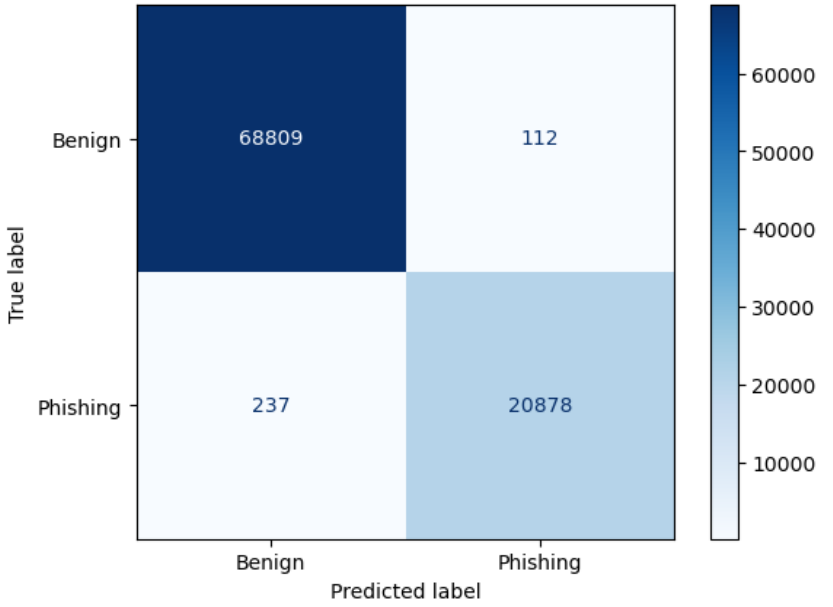

Bi-LSTM (99.02%)




GRU (99.60%)



# Load Bearing Test Results

| Model | Requests | Failures | Median Response Time (ms) | Avg Response Time(ms) | Min Response Time (ms) | Max Response Time (ms) | Requests/s | Failures/s |
|---|---|---|---|---|---|---|---|---|
| GRU | 35175 | 0 | 2100 | 2101.109 | 2051.363 | 2392.234 | 19.54203 | 0 |
| LSTM | 35233 | 0 | 2100 | 2105.836 | 2051.345 | 2360.184 | 19.57381 | 0 |
| BiLSTM | 35117 | 0 | 2100 | 2113.228 | 2051.292 | 2431.071 | 19.51131 | 0 |

| Model | Requests | Failures | Median Response Time (ms) | Avg Response Time(ms) | Min Response Time (ms) | Max Response Time (ms) | Requests/s | Failures/s |
|---|---|---|---|---|---|---|---|---|
| GRU | 34289 | 12577 | 5900 | 5702.544 | 3922.1665 | 9499.34 | 57.134340 | 20.956534 |
| LSTM | 34315 | 12845 | 6000 | 5698.6817 | 3357.4351 | 8613.556 | 57.174445 | 21.401887 |
| BiLSTM | 33816 | 12473 | 6000 | 5811.4993 | 4064.4171 | 13439.819 | 56.344 | 20.7825 |

# Results

- Used fewer testing epochs (10 instead of 50) due to the limitations of our hardware
- Training and Validation charts displayed the same trend with the relative accuracies also being similar
- Bi-LSTM –having the highest accuracy is shown as the lowest due to its hardware intensive labor (not able to test to full extent)
- GRU proves most effective for integration as it balances latency and accuracy

# Broader Impact

Although our work was successful, there are many areas with room for improvement. One of our main limitations in this projects includes access to powerful hardware causing time and computational constraints. Future work for this research may include:

- Multimodal implementation
- More involvement with email security
- Combining classification methods
- Training with additional, more advanced datasets

# References

Roy, S.S.; Awad, A.I.; Amare, L.A.; Erkihun, M.T.; Anas, M. Multimodel Phishing URL Detection Using LSTM, Bidirectional LSTM, and GRU Models. Future Internet 2022, 14, 340. https://doi.org/10.3390/fi14110340

Beasley, T., Blakely, K., & Tieman, M. (2024). Replication and extension of advanced phishing detection: Integrating machine learning with email services for malicious URL identification. CSCE 585 Project Report, University of South Carolina.