Creating a URL Phishing Detector

by Aggrey Timbwa, Richard Macharia, Pamela Jepkorir Chebii, Cynthia Njambi, Omara Waldea

Understanding Phishing Attacks

1 Business Risks

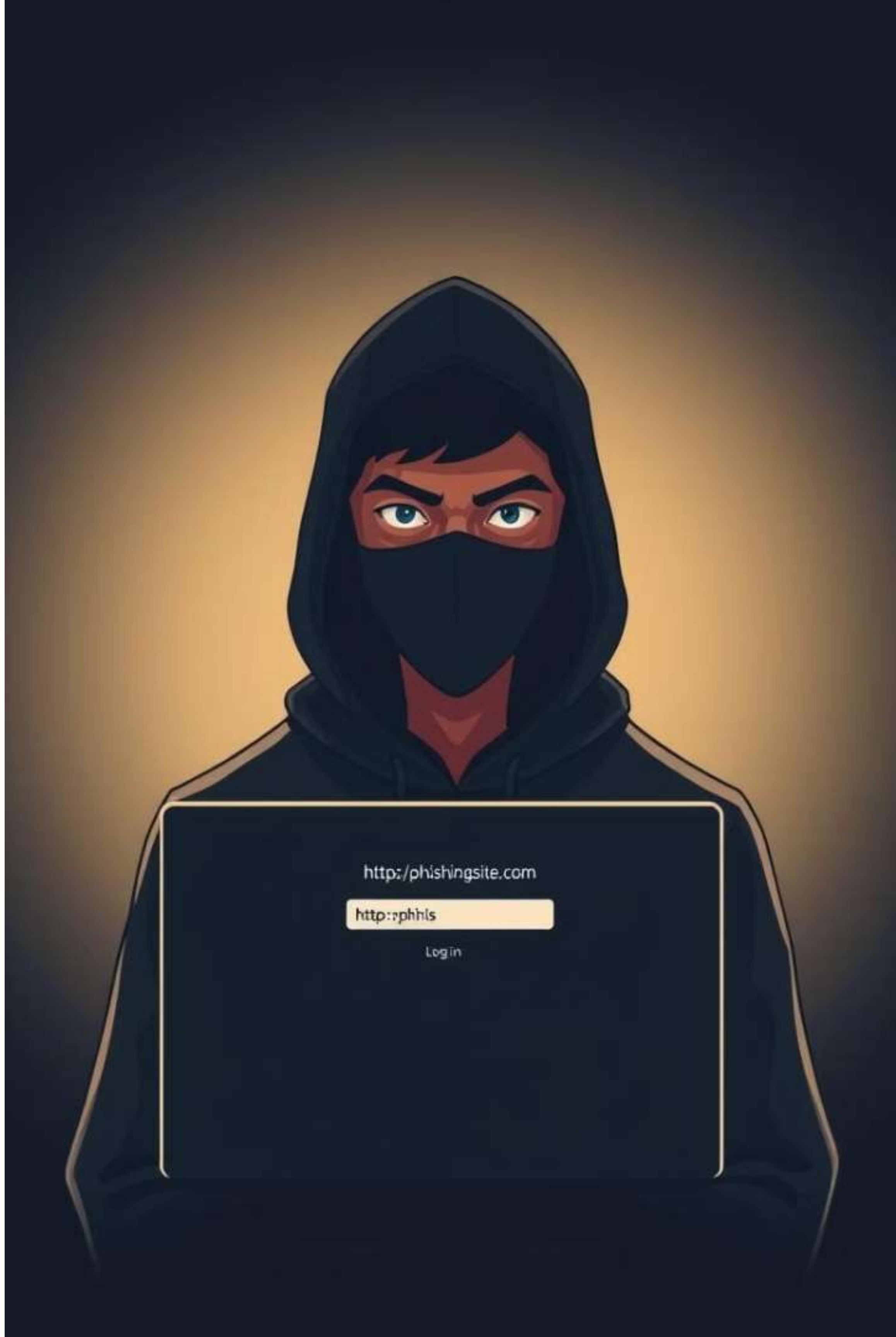
Financial losses and reputational damage.

2 Cybercrime Sophistication

Advanced tactics and social engineering.

3 Protection Necessity

Tools to identify phishing URLs for real-time detection.



Business Objectives

✓ High-Accuracy Detection

User-Friendly Deployment

Real-Time Classification

Preparing a Training Dataset

1

Data Collection

Gather diverse URLs from multiple sources.

2

Data Preprocessing

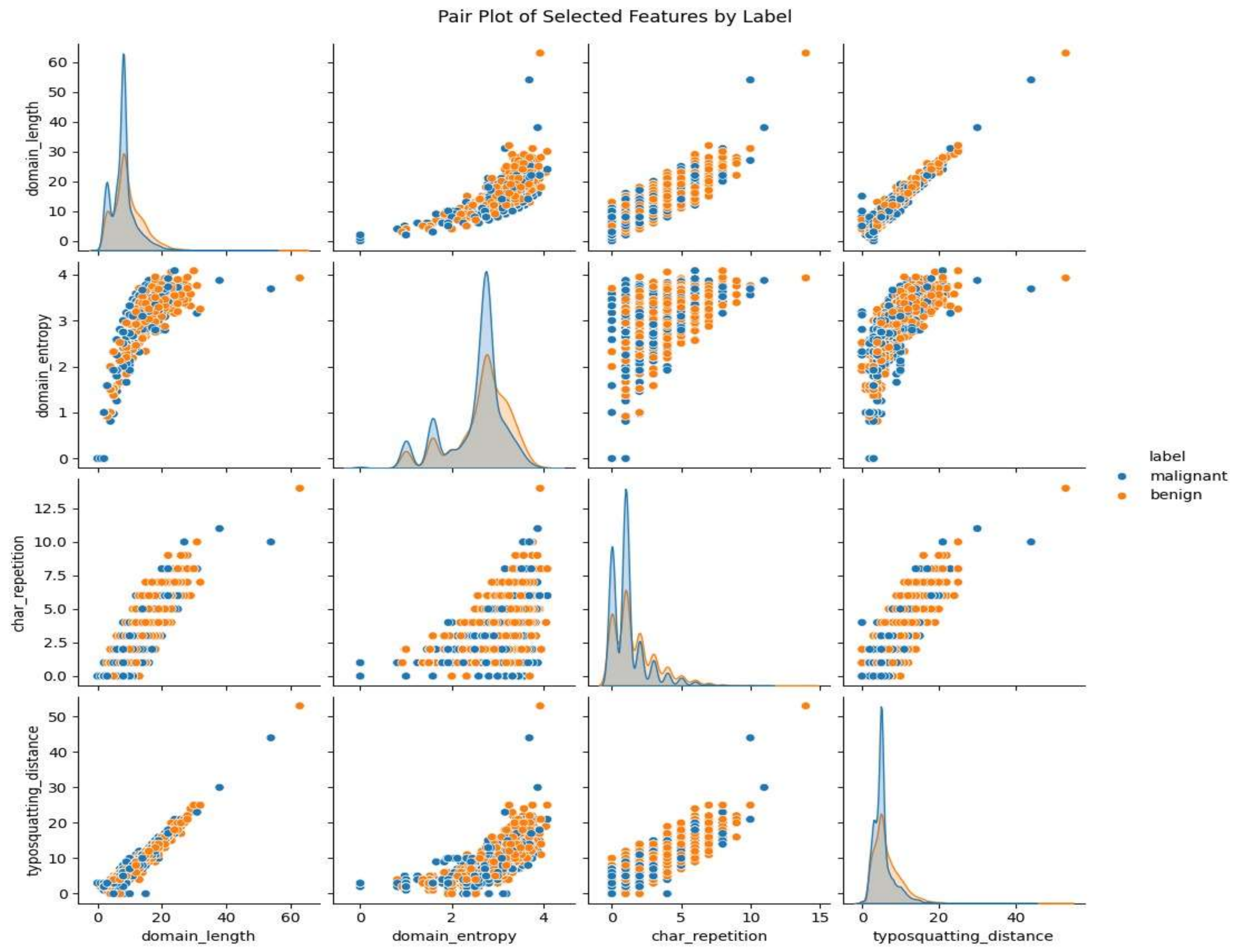
Clean dataset by removing duplicates and invalid entries.

3

Feature Extraction and Engineering

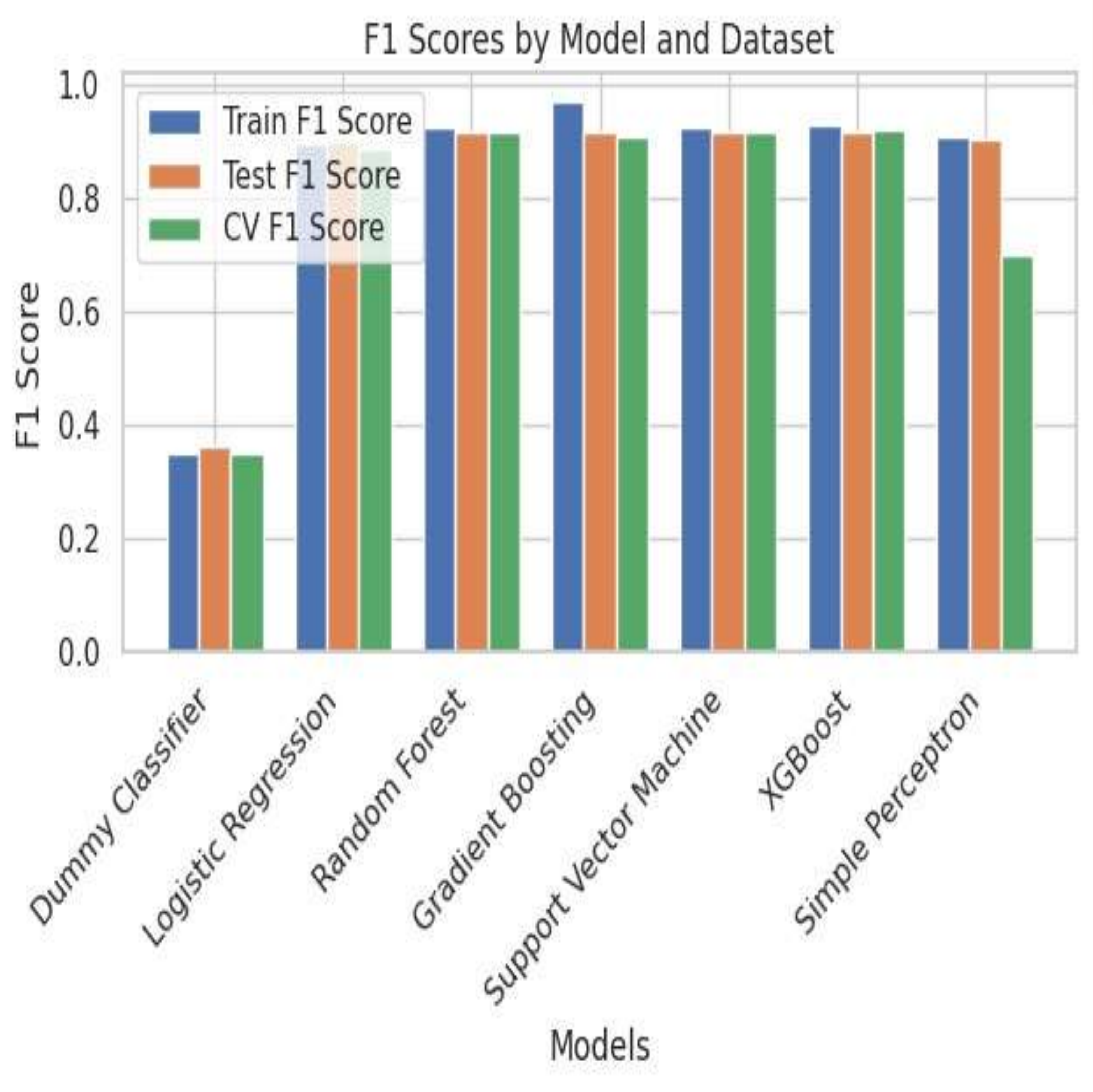
Extract relevant URL attributes like domain age.

.



The pair plot shows distinct feature distributions and relationships between benign and malignant labels, with noticeable patterns in domain length, entropy, character repetition, and typosquatting distance, indicating potential separation between the two classes based on these features.

Model Training and Evaluation

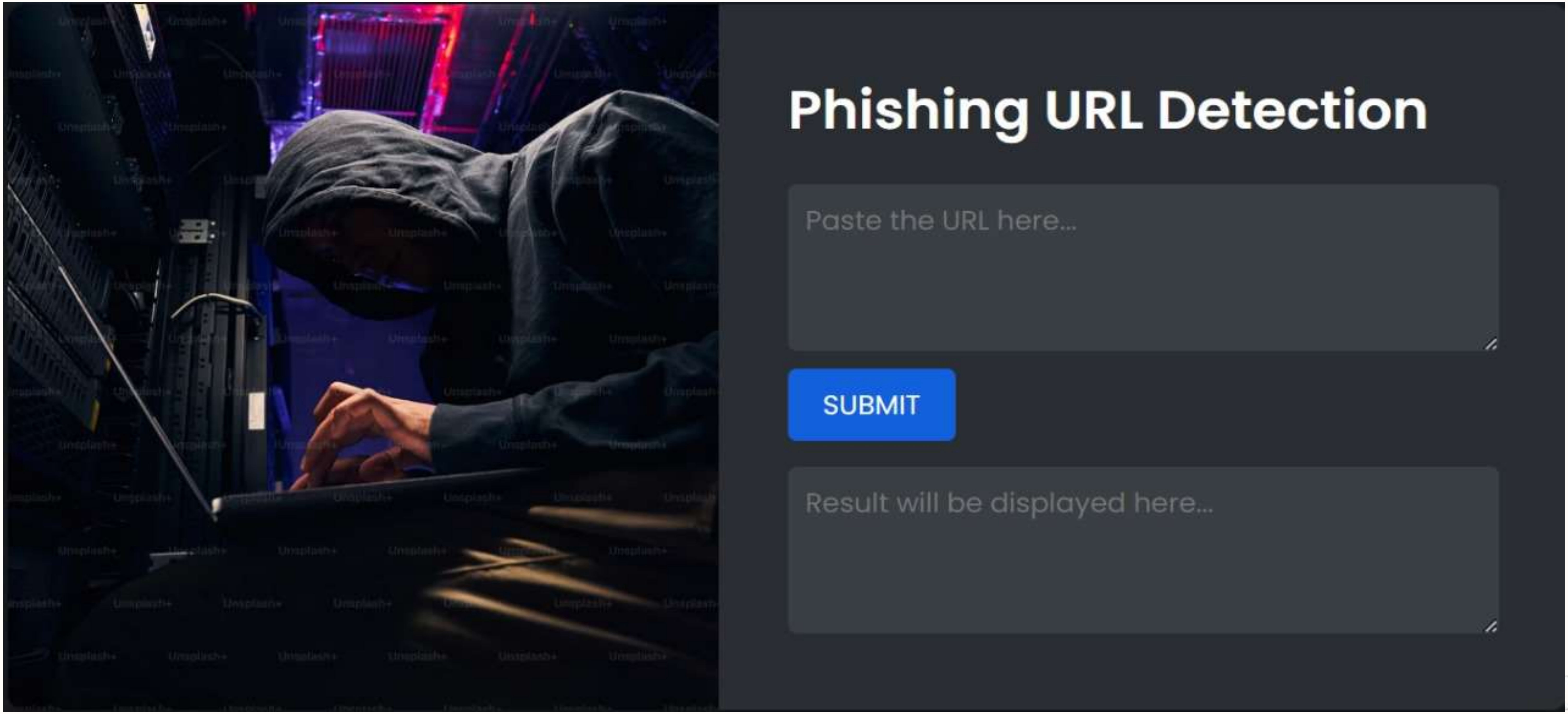


The XGBoost and Gradient Boosting models exhibit the highest performance across all metrics, while Random Forest and Support Vector Machine also perform well;

Logistic Regression and Simple Perceptron demonstrate good but slightly lower performance

The Dummy Classifier shows the lowest scores, indicating it is a poor benchmark.

Deploying the Detector



Phishing URL Detection

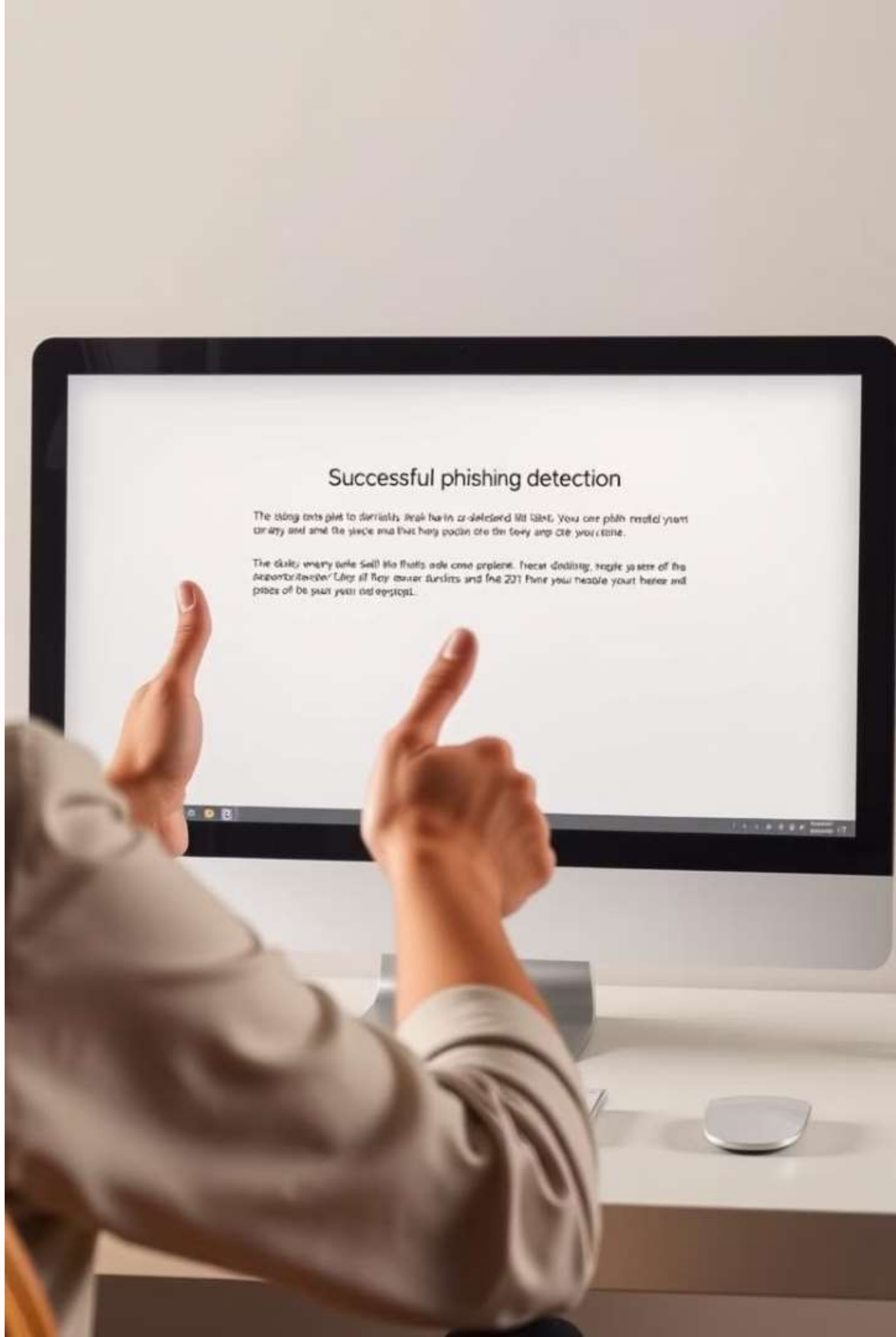
Paste the URL here...

Result will be displayed here...

Conclusion

High-accuracy phishing detection can be achieved with models like XGBoost and Gradient Boosting

A URL phishing detector is an effective tool for protecting individuals and organizations from phishing attacks.



Q&A

Thank You