Privacy And Security With Deep Learning

Phishing URL Detection With Deep Learning

HOCHSCHULE
HAMM-LIPPSTADT

Name : MD Sayem
2200025

Privacy And Security

- Data Confidentiality
- Data Integrity
- Access Control

Cyber Attacks

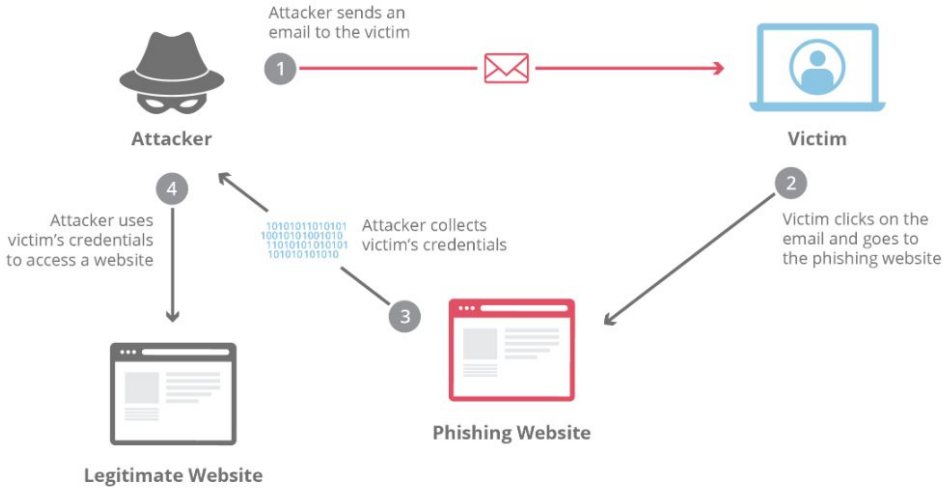
An unauthorized attempt to access a system with the intent to steal sensitive information or cause harm.

Types OF Cyber Attacks

- Malware
- Denial of Service (DoS) Attack
- Spoofing
- Man-in-the-Middle (MitM) Attack
- Password Attack
- Ransomware
- **Phishing Attack**

Phishing Attack

Creating fake environments to trick people into revealing sensitive information like passwords or credit card details.



Types Of Phishing

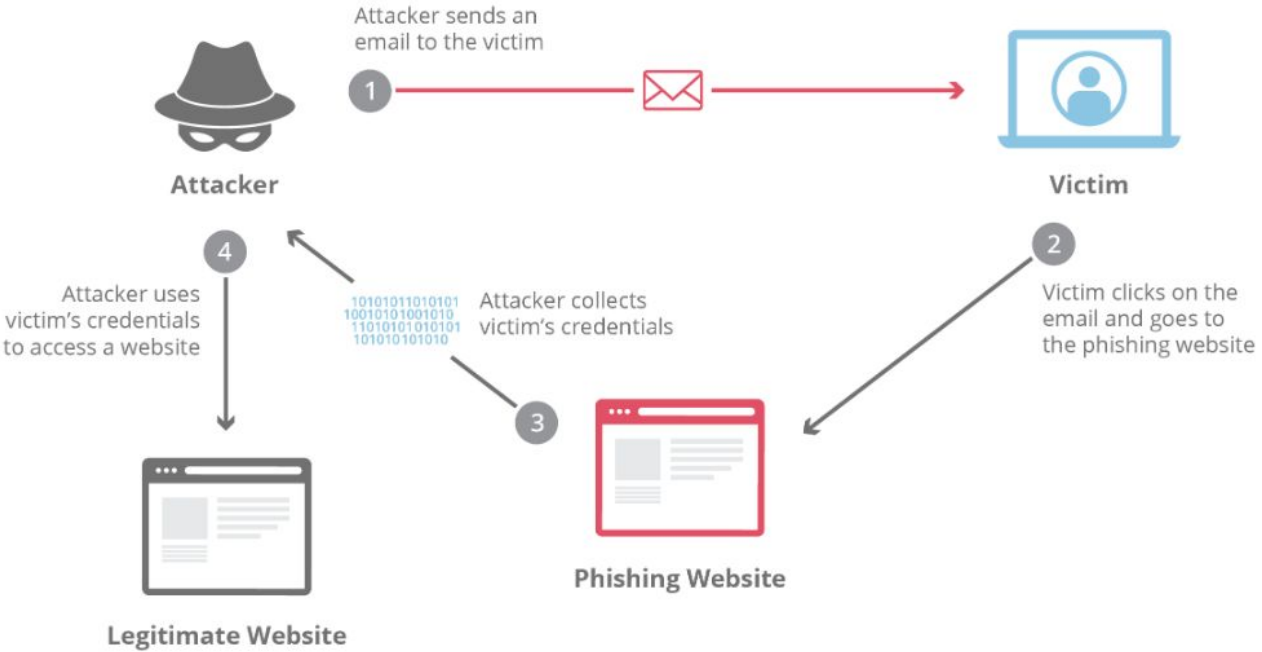
- Email Phishing
- Spear Phishing
- Smishing (SMS Phishing)
- Vishing (Voice Phishing)
- **URL Based Phishing**

URL Based Phishing

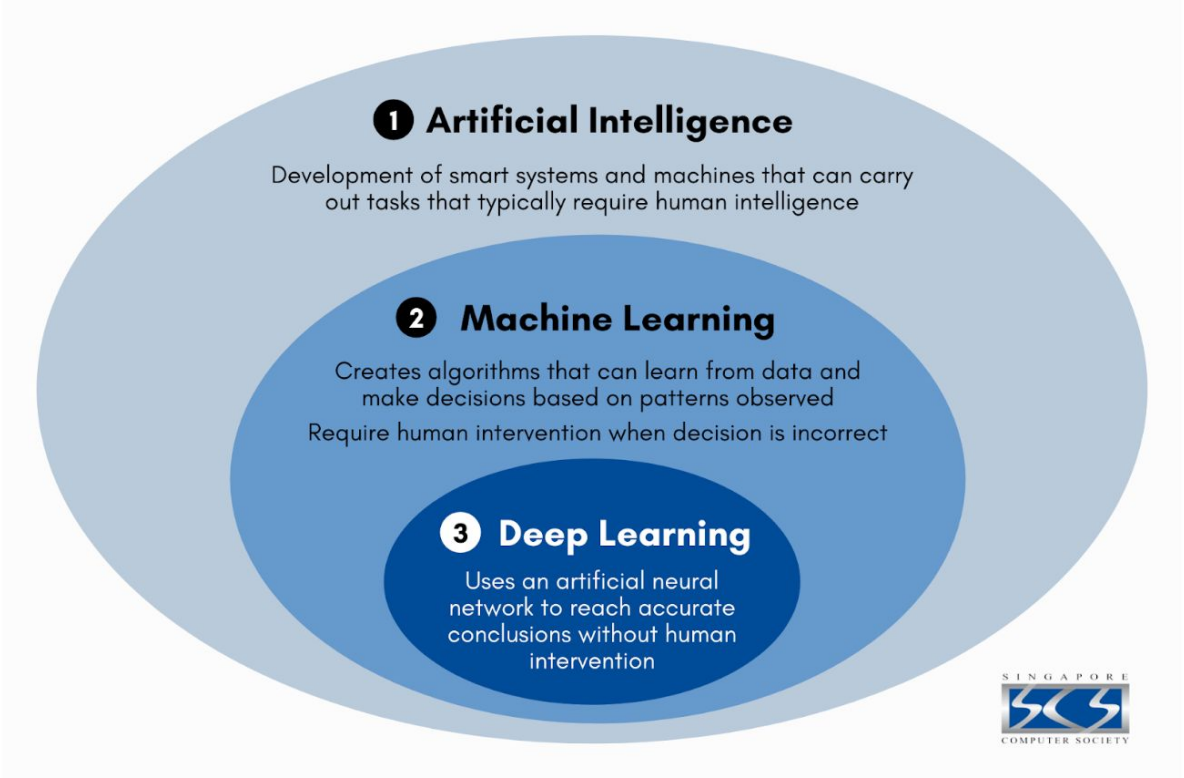
Attackers use deceptive URLs to trick users into visiting malicious websites, designed to steal sensitive information like usernames, passwords, or financial data.

https://paypa1-login.com
paypa1.com
g00gle.com
https://faceb00k-security.com

Phishing Operation



Deep Learning



Dataset

SN	Feature	Description	Type
F0	Type	Indicating the type of the URL. It is a Boolean feature with 0 representing a legitimate URL and 1 representing a phishing URL.	Boolean
F1	url_length	Representing the number of characters in a URL, including the domain name, path, and any query parameters.	Numeric
F2	number_of_dots_in_url	Indicating the number of dots (".") in the UR	Numeric
F3	having_repeated_digits_in_url	A Boolean feature that denotes whether the URL has repeated digits (e.g., 2232)	Boolean
F4	number_of_digits_in_url	Representing the number of digits (0-9) in the URL.	Numeric
F5	number_of_special_char_in_url	Indicating the number of special characters (e.g., ", #, $, %, &, ~) in the URL.	Numeric
F6	number_of_hyphens_in_url	Representing the number of hyphens ("-") in the URL.	Numeric
F7	number_of_underline_in_url	Indicating the number of underscores ("_") in the URL.	Numeric
F8	number_of_slash_in_url	Representing the number of forward slashes ("/") or backward slashes ("\") in the URL.	Numeric
F9	number_of_questionmark_in_url	Indicating the number of question marks ("?") in the URL.	Numeric
F10	number_of_equal_in_url	Representing the number of equal signs ("=") in the URL. It is a numeric feature	Numeric
F11	number_of_at_in_url	Indicating the number of at symbols ("@") in the URL.	Numeric
F12	number_of_dollar_sign_in_url	Representing the number of dollar signs ("$") in the URL.	Numeric
F13	number_of_exclamation_in_url	Indicating the number of exclamation marks ("!") in the URL.	Numeric
F14	number_of_hashtag_in_url	Representing the number of hashtags ("#") in the URL.	Numeric
F15	number_of_percent_in_url	Indicating the number of percent signs (%) in the URL.	Numeric

- 247950 instances,
- 128541 Phishing URLs
- 119409 legitimate URLs.
- 41 features
- 1 target variable (0=legitimate,1=phishing)

Dataset link : https://data.mendeley.com/datasets/6tm2d6sz7p/1

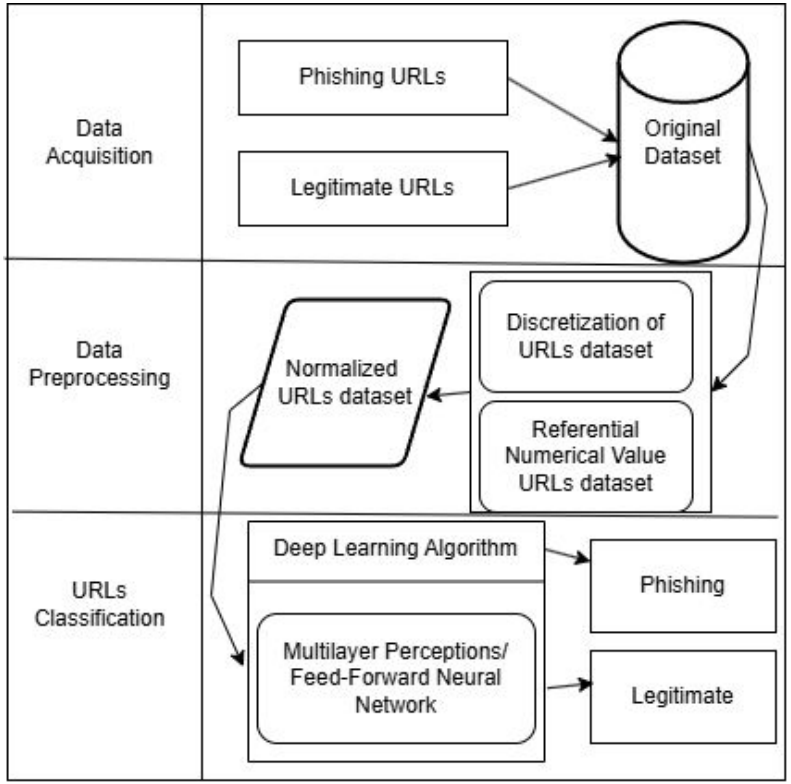
Phishing URLs And DL

Typical Patterns/Features in URL
(e.g., length, special characters, number of dots, number of digits, number of hyphens etc.).

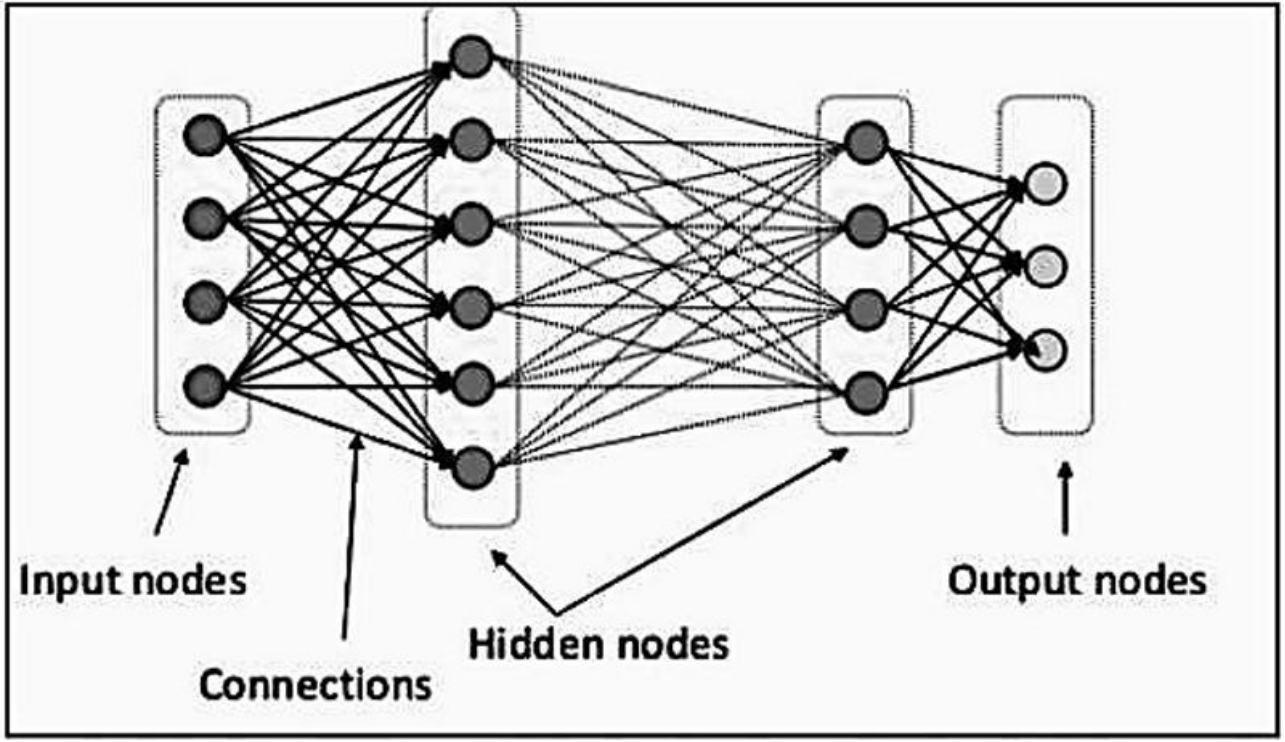
Input: Features of the URL

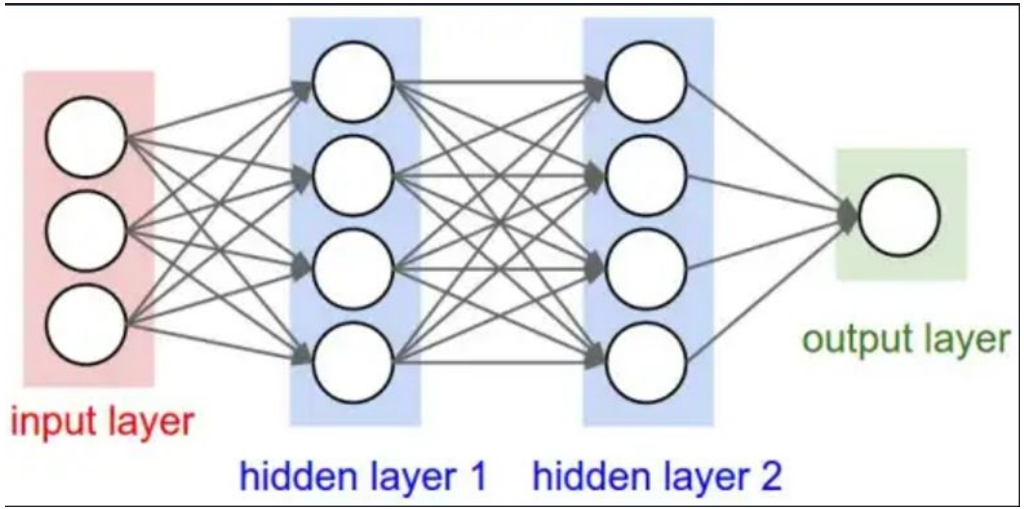
Hidden Layers: The model identifies patterns in the URL features

Output: Predicts whether a URL is phishing or legitimate.



Feed Forward Neural Network (FNN)/MLP (Multi-Layer Perceptron)





	Type	url_length	number_of_dots_in_url	having_repeated_digits_in_url	number_of_digits_in_url	number_of_special_char_in_url	num
0	0	37	2	0	0	8	
1	1	70	5	0	0	12	
2	0	42	2	0	6	8	
3	0	46	2	0	0	7	
4	0	51	3	0	0	9	
...	...	...	...	...	...	...	
247945	0	42	1	0	0	6	
247946	0	42	2	0	0	8	
247947	1	33	2	0	0	8	
247948	1	83	1	1	19	9	
247949	0	34	3	0	0	7	

247950 rows × 42 columns

Data Splitting

Training Set Size:	198360 samples 80%
Validation Set Size:	24795 samples 10%
Testing Set Size:	24795 samples 10%
Input shape	41

```
data=pd.read_csv('Dataset.csv')
total_nulls = data.isnull().sum().sum()
print(f"There are {total_nulls} null values in the dataset.")
```

```
There are 0 null values in the dataset.
```

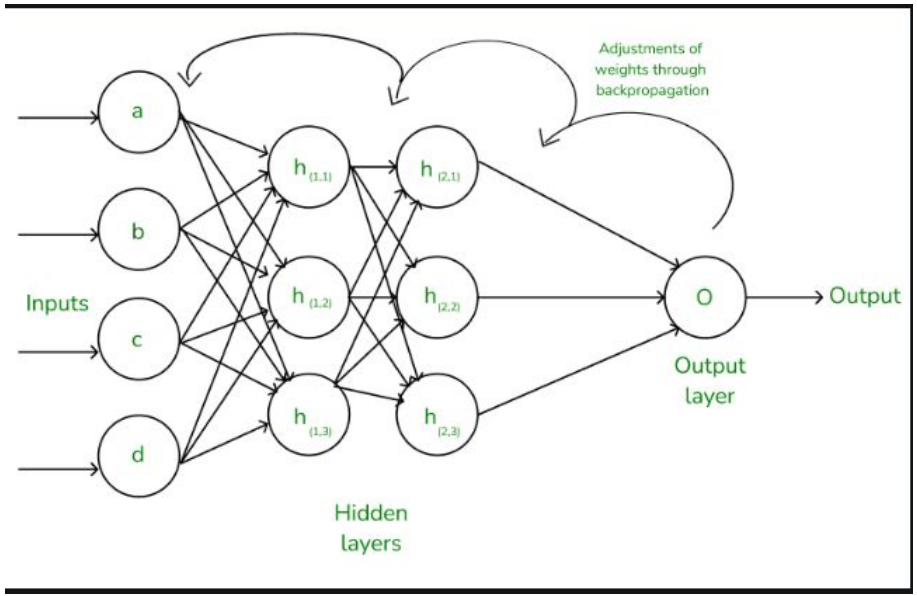
DL Algorithm

```python
from tensorflow import keras
from tensorflow.keras import layers

# Building the neural network model
model = keras.Sequential([
    layers.BatchNormalization(input_shape=input_shape),
    layers.Dense(512, activation='relu'),
    layers.BatchNormalization(),
    layers.Dropout(0.3),
    layers.Dense(512, activation='relu'),
    layers.BatchNormalization(),
    layers.Dropout(0.3),
    layers.Dense(1, activation='sigmoid'),  # Binary classification output
])
```

```python
# Compile the model
model.compile(
    optimizer='adam',
    loss='binary_crossentropy',
    metrics=['binary_accuracy', keras.metrics.Precision(), keras.metrics.Recall()],
)
```

Early Stopping And Backpropagation

```
# Early stopping to avoid overfitting
early_stopping = keras.callbacks.EarlyStopping(
    patience=20,
    min_delta=0.01,
    restore_best_weights=True,
)

# Train the model
history = model.fit(
    X_train, y_train,
    validation_data=(X_valid, y_valid),
    batch_size=512,
    epochs=200,
    callbacks=[early_stopping],
)
```



Results

Loss: Measures how well the model predicts the data.

Accuracy: The overall correctness of predictions.

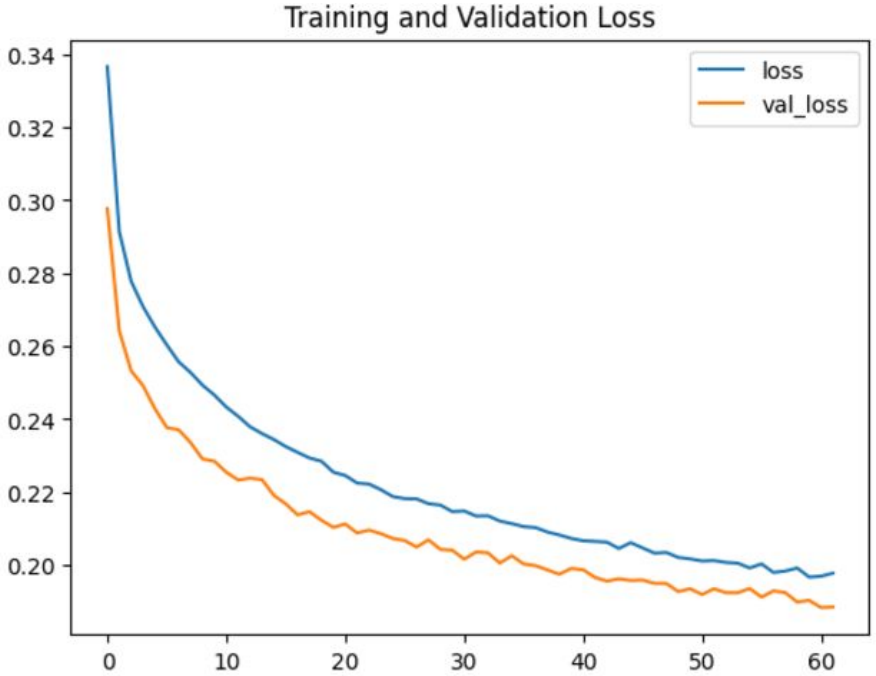
--- Best Results ---

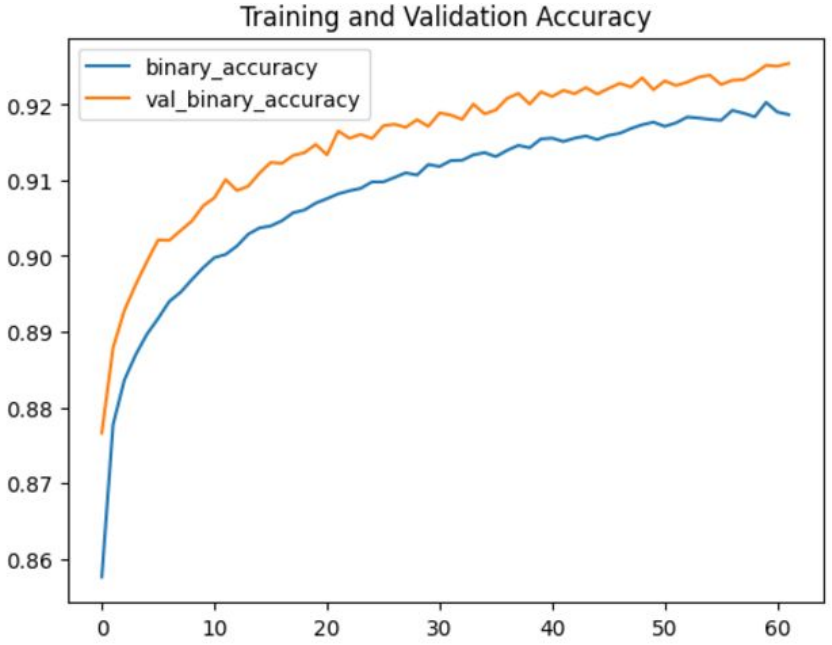
Metric	Value
Best Validation Loss	0.18840599060058594
Best Validation Accuracy	0.9254285097122192
Best Recall	0.9025207161903381
Best Precision	0.9493273496627808

Recall: How many of the actual phishing URLs were correctly detected.

Precision: How many of the predicted phishing URLs were correct.

Results


Training and Validation Loss
loss
val_loss
0.34
0.32
0.30
0.28
0.26
0.24
0.22
0.20
0
10
20
30
40
50
60


Training and Validation Accuracy
binary_accuracy
val_binary_accuracy
0.92
0.91
0.90
0.89
0.88
0.87
0.86
0
10
20
30
40
50
60

Results

Loss: Measures how well the model predicts the data.

Accuracy: The overall correctness of predictions.

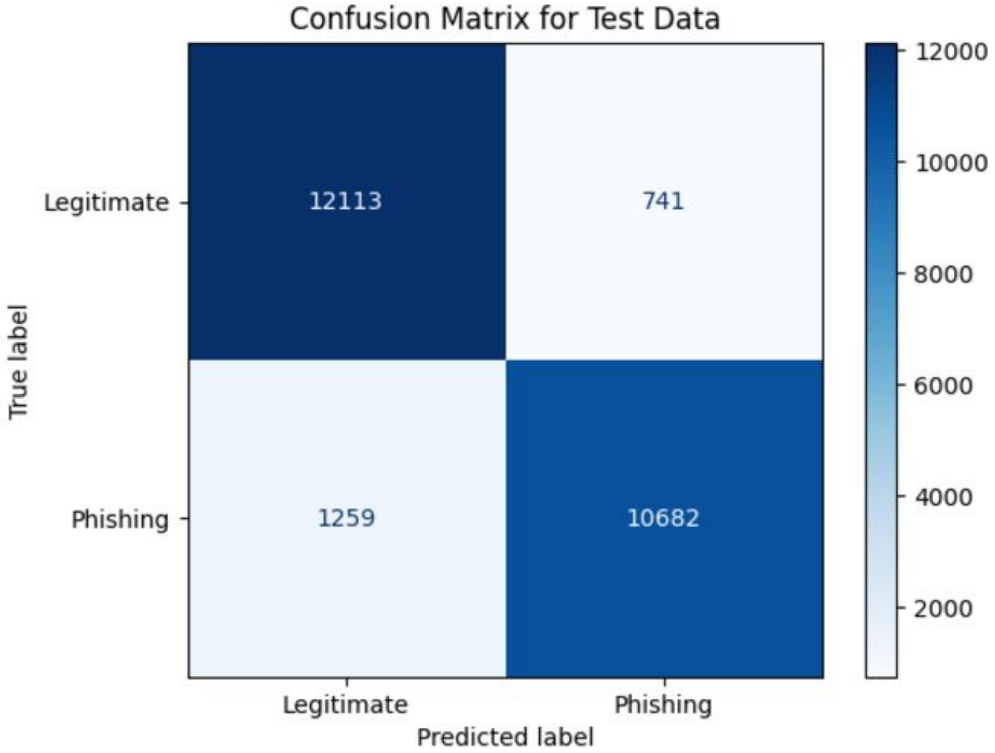
--- Test Results ---

Metric	Value
Test Loss	0.196093276143074 04
Test Accuracy	0.919338583946228
Test Precision	0.9351308941841125
Test Recall	0.8945649266242981

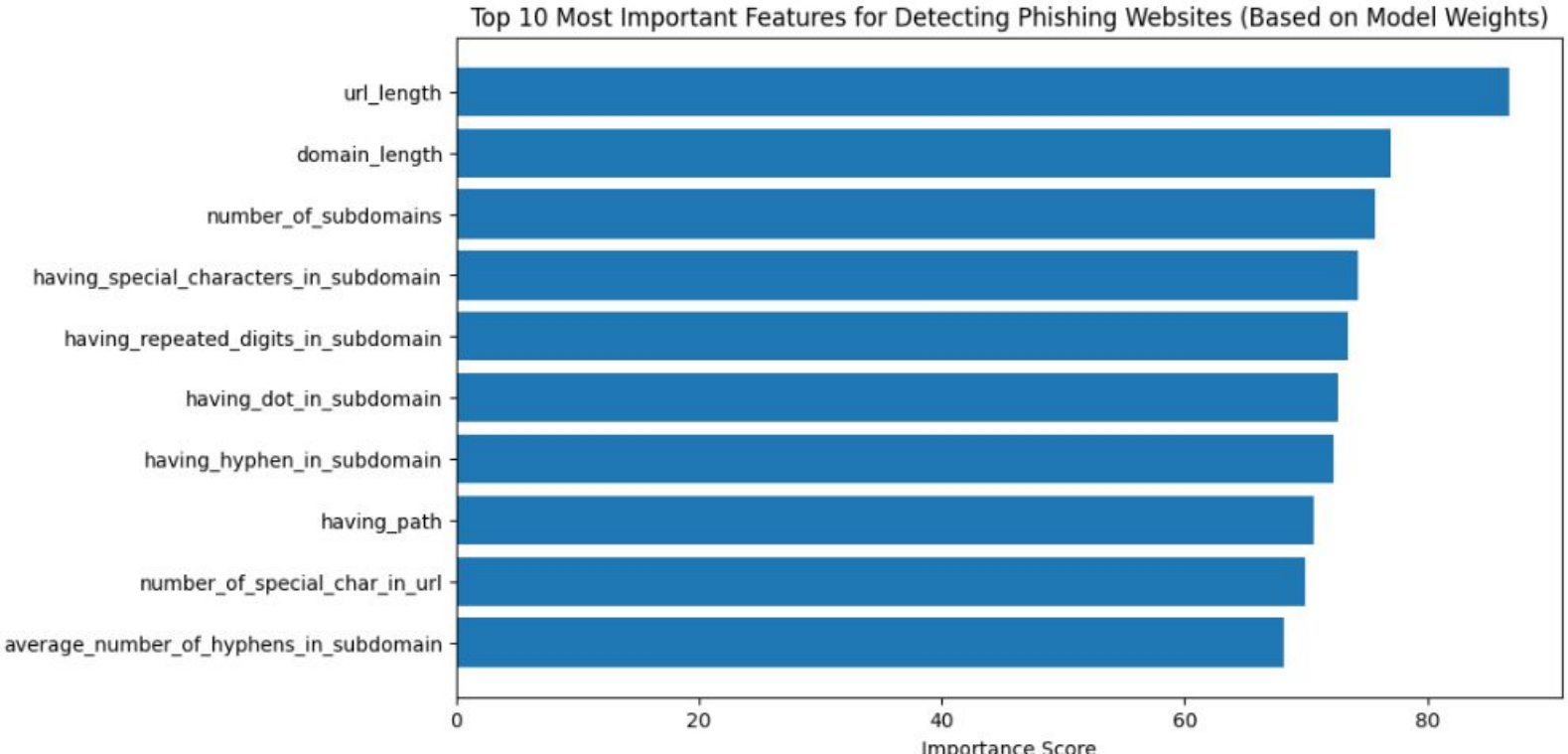
Recall: How many of the actual phishing URLs were correctly detected.

Precision: How many of the predicted phishing URLs were correct.

Results


Confusion Matrix for Test Data
Legitimate
Phishing
True label
12113
741
1259
10682
Legitimate
Phishing
Predicted label
12000
10000
8000
6000
4000
2000

Results


Top 10 Most Important Features for Detecting Phishing Websites (Based on Model Weights)
url_length
domain_length
number_of_subdomains
having_special_characters_in_subdomain
having_repeated_digits_in_subdomain
having_dot_in_subdomain
having_hyphen_in_subdomain
having_path
number_of_special_char_in_url
average_number_of_hyphens_in_subdomain
0
20
40
60
80
Importance Score

Thank You

Question ?