SmartInternz + IBM

Web Phishing Detection Using Deep Learning

INTELLIGENT SYSTEMS SQUAD

Team Members

- Ambuj Gupta (19BCE0366)
- Karan Rochlani (19BCE0383)
- Harshal Vijay Ghadge (19BAI10033)
- DVS Hitesh Reddy (19BCE1066)

Introduction

- There are a number of users who purchase products online and make payments through e-banking. There are e-banking websites that ask users to provide sensitive data such as username, password & credit card details, etc., often for malicious reasons.
- This type of e-banking website is known as a phishing website. Web service is one of the key communications software services for the Internet.
- Web phishing is one of many security threats to web services on the Internet.
- Common threats of web phishing:
- Web phishing aims to steal private information, such as usernames, passwords, and credit card details, by way of impersonating a legitimate entity.
- It will lead to information disclosure and property damage.
- Large organizations may get trapped in different kinds of scams.

“To Detect Web Phishing Using Deep Learning Techniques in IBM Watson”

AIM OF OUR PROJECT

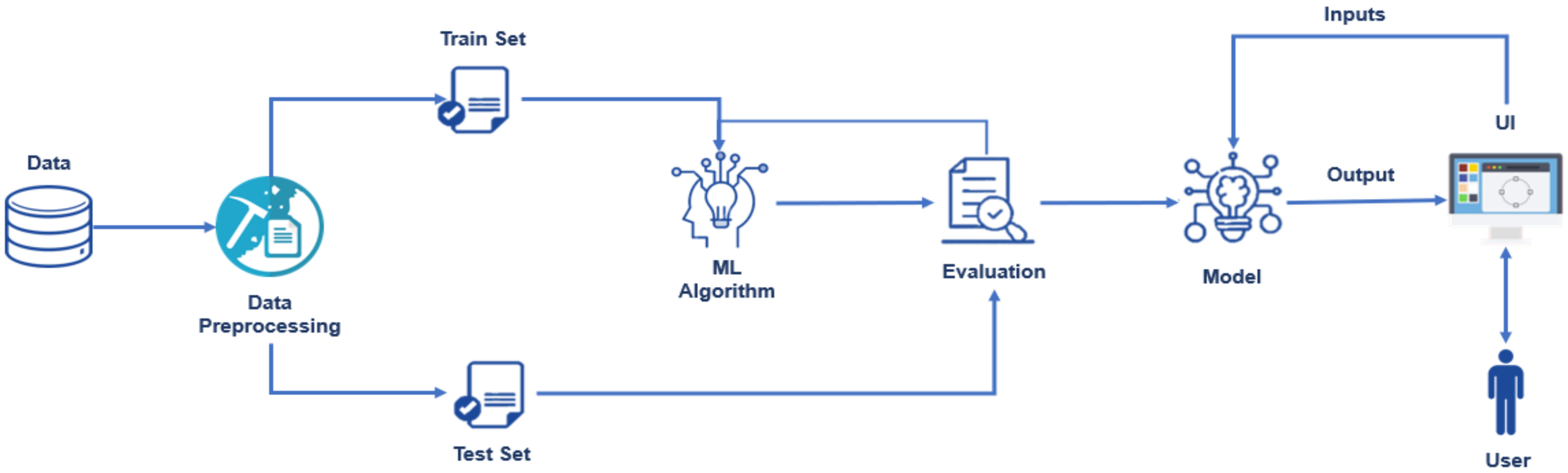
Problem Statement
Web Phishing Detection

- You'll be able to understand the problem to classify if it is a regression or a classification kind of problem.
- You will be able to know how to pre-process/clean the data using different data pre-processing techniques.

Solution

- Using Artificial Neural Network Model (ANN) this project is developed.Train the machine with preprocessed data using an appropriate machine learning algorithm.
- Save the model and its dependencies.

Architecture Diagram



Project Flow

- Installing the required packages and libraries.
- Importing the required libraries for the model to run.
- Download the dataset,feeding it to the model, and understanding the dataset
- Data Preprocessing – Checking for outliers and null values. If there any null values we use Label Encoding to convert then into binary format.
- Dividing the model into Train and Test data. Fitting the model and predicting.
- Building Flask Web Application

Experimental Analysis

Model Structure

```
In [121]: from tensorflow.keras.models import Sequential
          from tensorflow.keras.layers import Dense,Dropout

In [122]: Regression_Model = Sequential()

In [123]: Regression_Model.add(Dense(units=30, kernel_initializer="random_uniform", activation="relu"))

In [124]: Regression_Model.add(Dense(units=60, kernel_initializer="random_uniform", activation="relu"))

In [125]: Regression_Model.add(Dropout(0.2))

In [126]: Regression_Model.add(Dense(units=60, kernel_initializer="random_uniform", activation="relu"))

In [127]: Regression_Model.add(Dropout(0.2))

In [128]: Regression_Model.add(Dense(units=1, kernel_initializer="random_uniform", activation="sigmoid"))

In [129]: Regression_Model.compile(optimizer = "rmsprop",loss ="binary_crossentropy", metrics = ["accuracy"] )

In [130]: history = Regression_Model.fit(x_train,y_train, batch_size =64,epochs = 100,validation_data=(x_test, y_test))
          uracy: 0.9653
          Epoch 95/100
          130/130 [==============================] - 0s 3ms/step - loss: 0.0404 - accuracy: 0.9846 - val_loss: 0.2413 - val_acc
          uracy: 0.9616
          Epoch 96/100
          130/130 [==============================] - 0s 3ms/step - loss: 0.0395 - accuracy: 0.9828 - val_loss: 0.2133 - val_acc
          uracy: 0.9649
          Epoch 97/100
          130/130 [==============================] - 0s 3ms/step - loss: 0.0381 - accuracy: 0.9840 - val_loss: 0.2342 - val_acc
          uracy: 0.9656
```

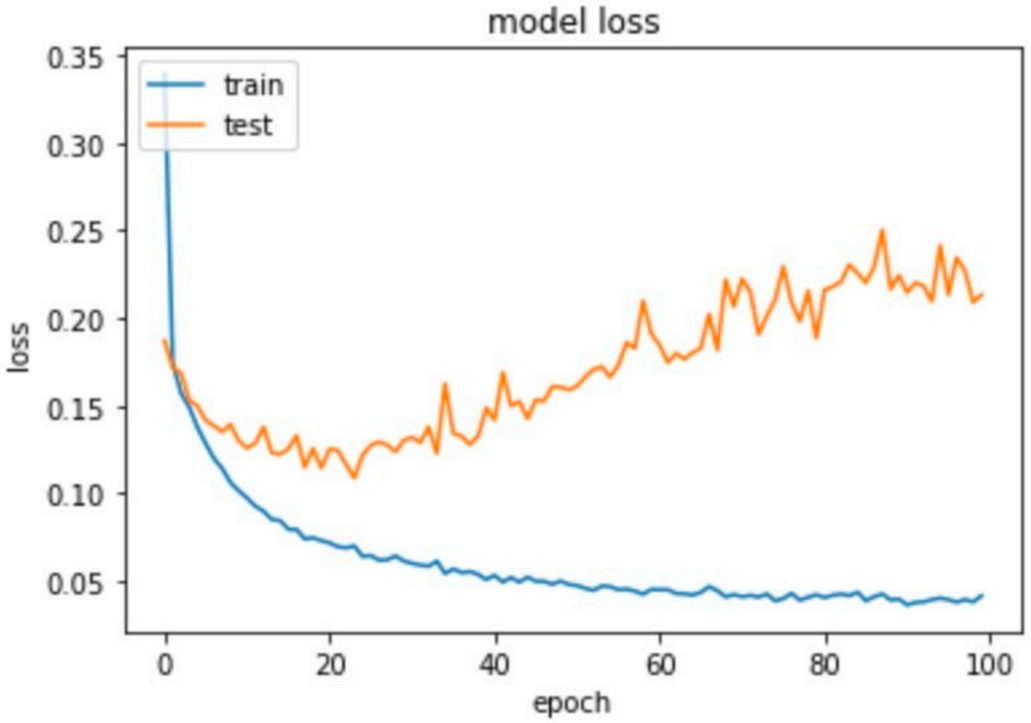
98.35%

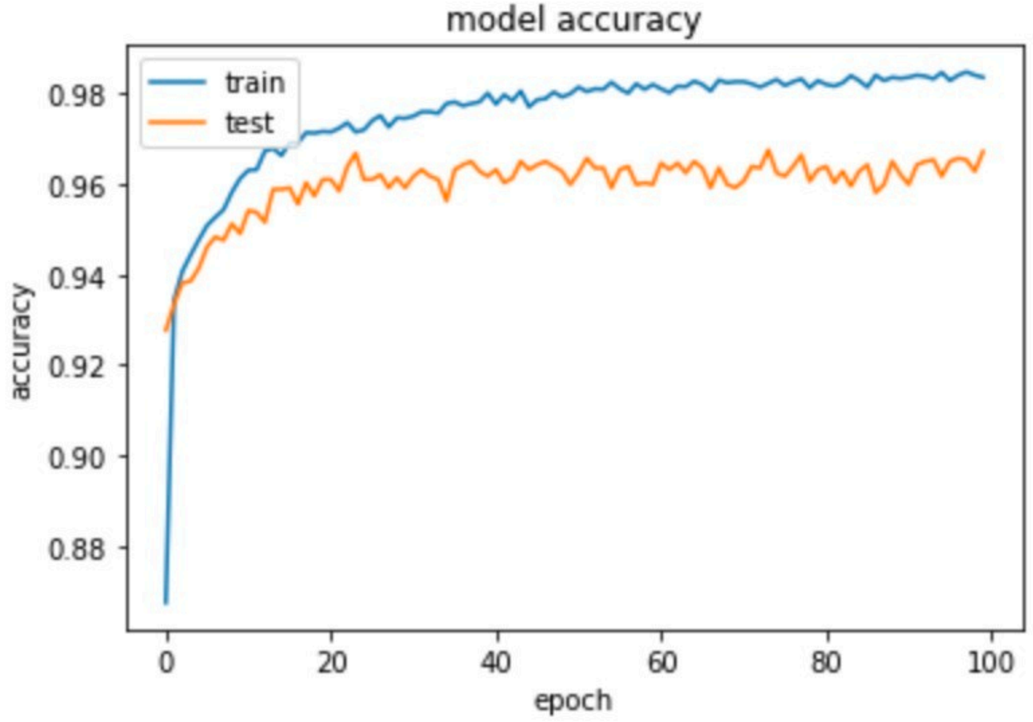
Training Accuracy

96.7%

Testing Accuracy

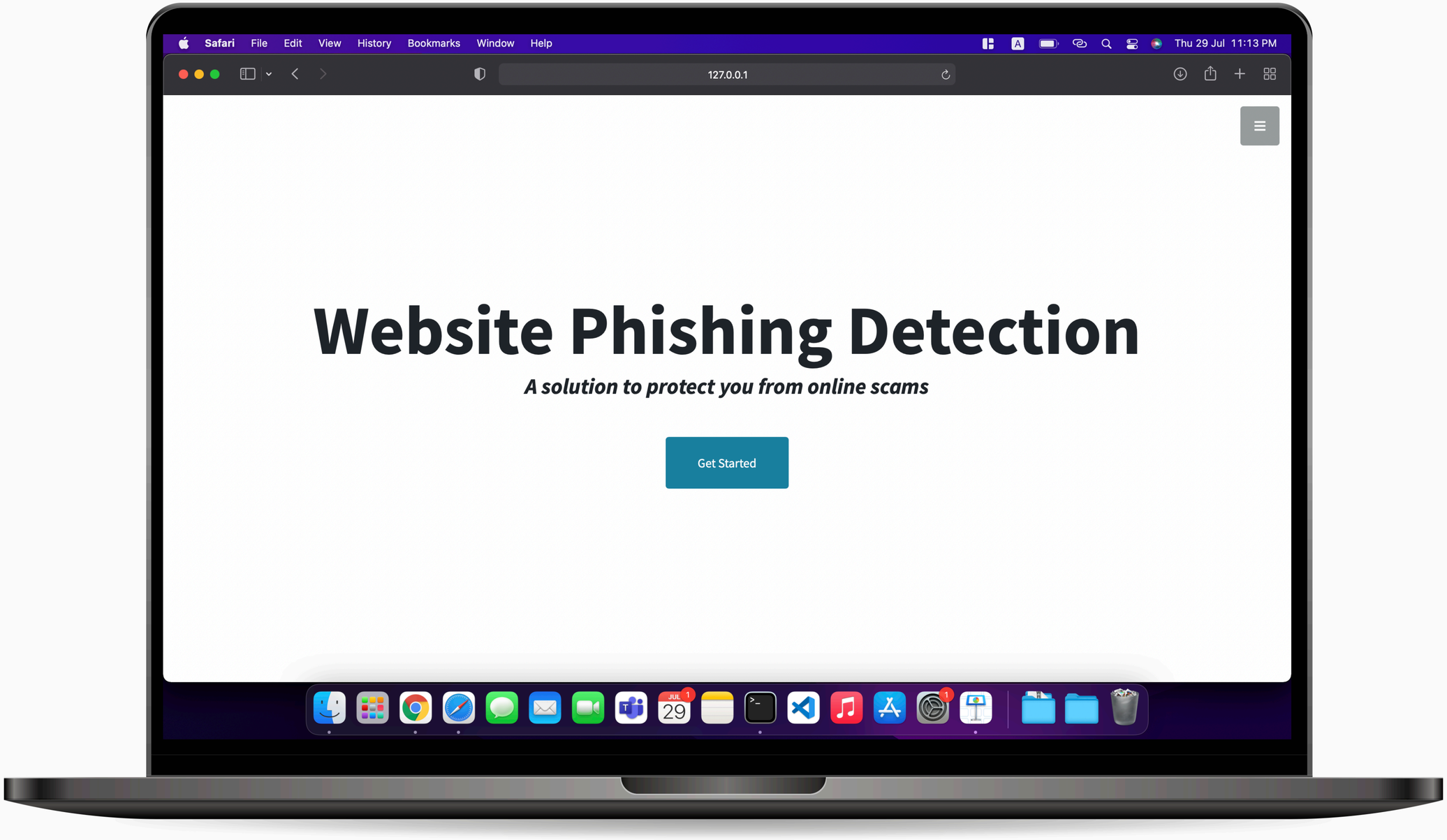
In [37]:
```
import matplotlib.pyplot as plt
plt.plot(history.history['accuracy'])
plt.title('model accuracy')
plt.ylabel('accuracy')
plt.xlabel('epoch')
plt.legend(['train', 'test'], loc='upper left')
plt.show()
# summarize history for loss
plt.plot(history.history['loss'])
plt.title('model loss')
plt.ylabel('loss')
plt.xlabel('epoch')
plt.legend(['train', 'test'], loc='upper left')
plt.show()
```

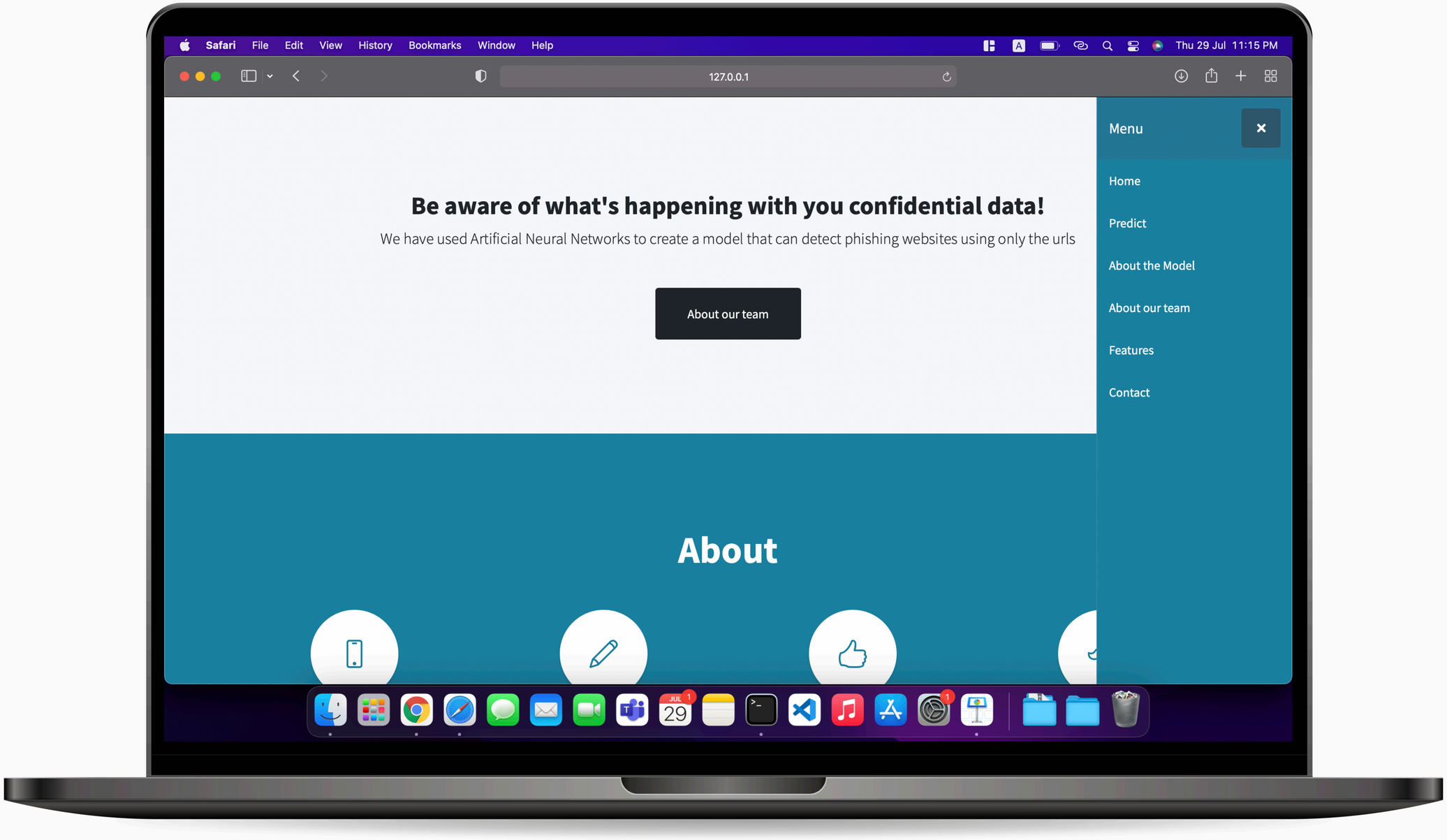




Web Design and UI

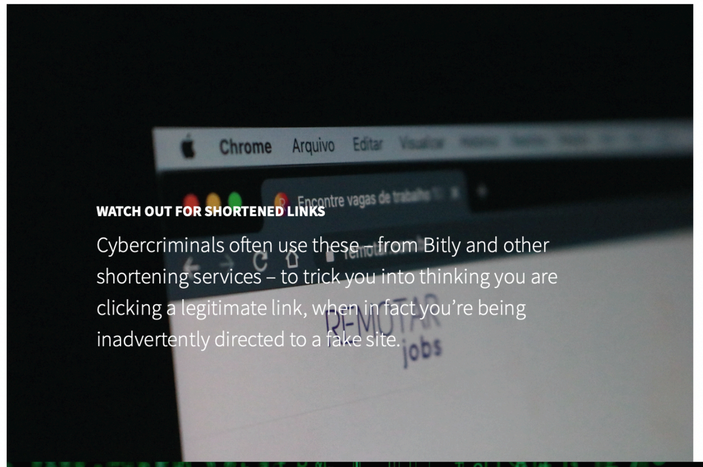
DEVELOPED USING BOOTSTRAP


Safari File Edit View History Bookmarks Window Help
Thu 29 Jul 11:13 PM
127.0.0.1
Website Phishing Detection
A solution to protect you from online scams
Get Started


Safari File Edit View History Bookmarks Window Help
Thu 29 Jul 11:15 PM
127.0.0.1
Be aware of what's happening with you confidential data!
We have used Artificial Neural Networks to create a model that can detect phishing websites using only the urls
About our team
Menu
Home
Predict
About the Model
About our team
Features
Contact
About

FEATURES

PROTECT YOURSELF FROM PHISHING ATTACKS

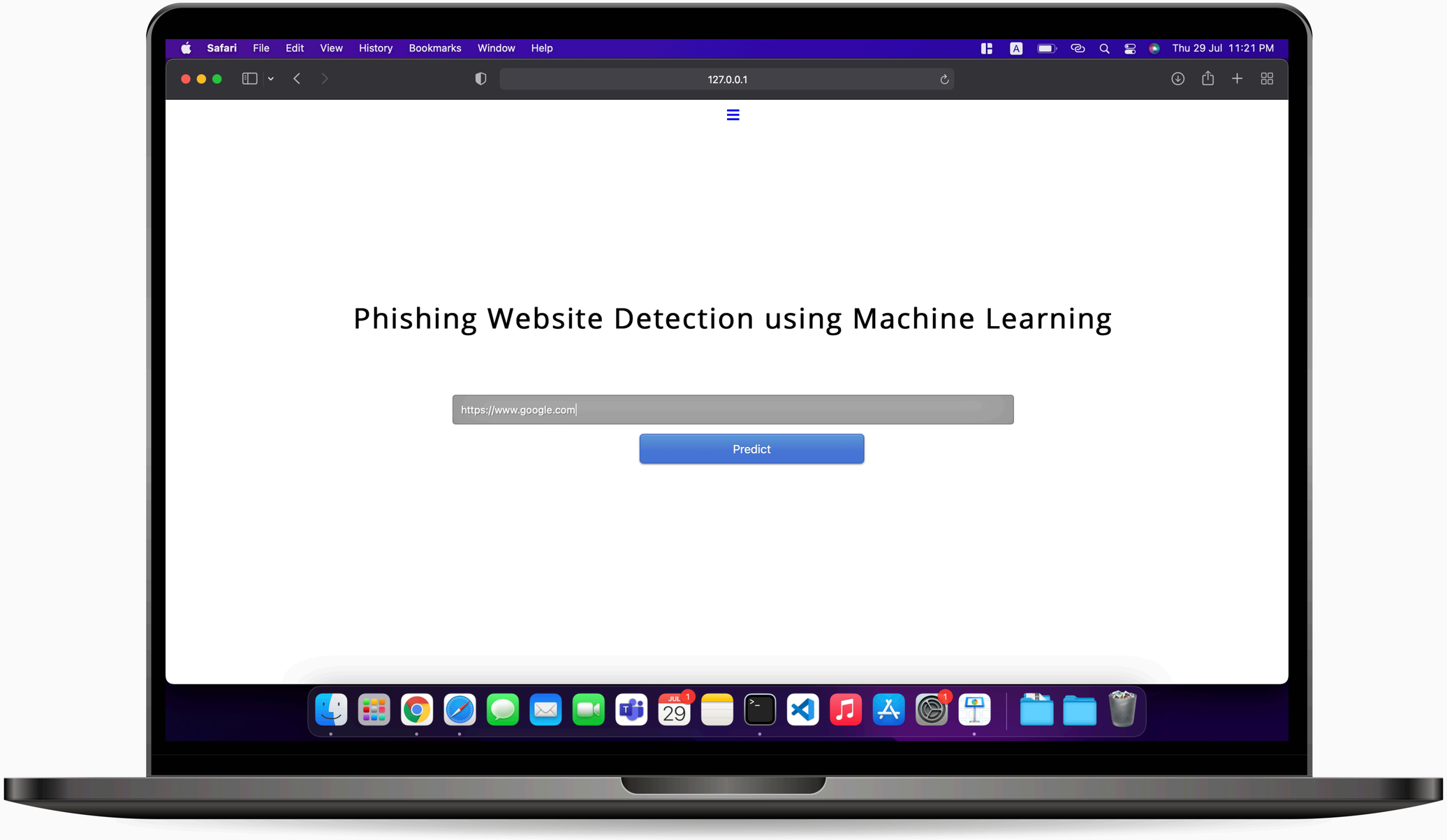


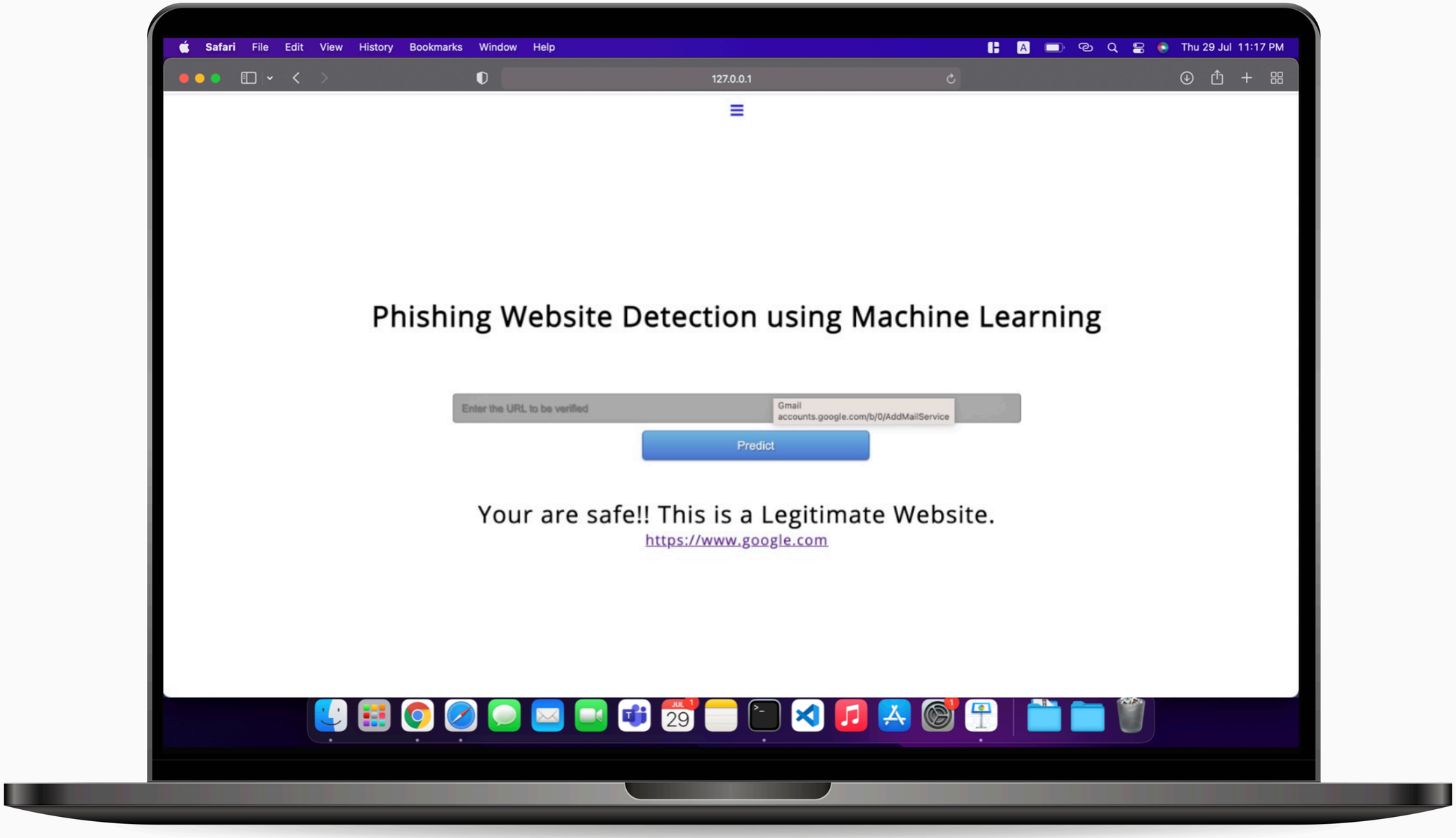
WATCH OUT FOR SHORTENED LINKS

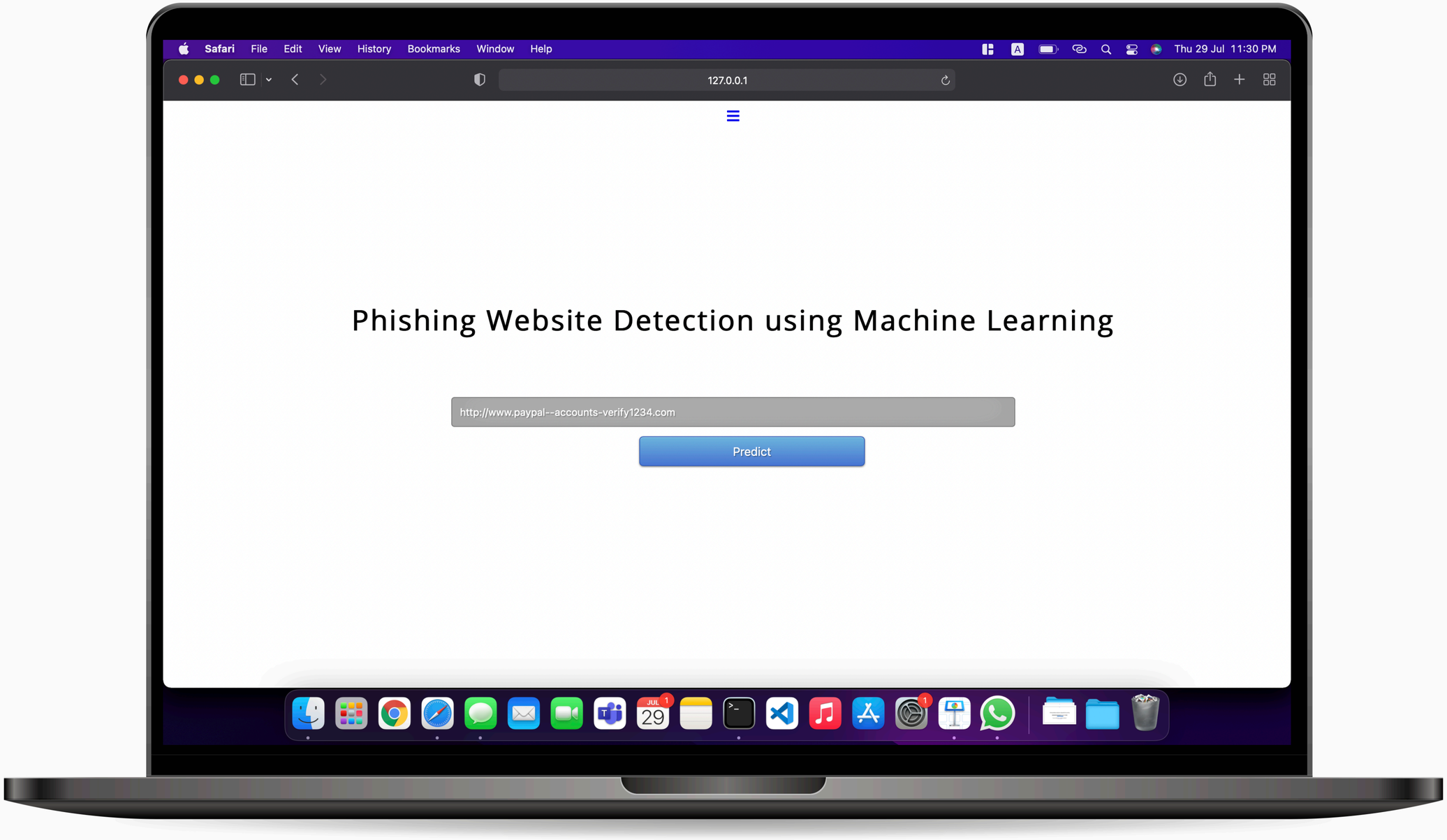
Cybercriminals often use these – from Bitly and other shortening services – to trick you into thinking you are clicking a legitimate link, when in fact you're being inadvertently directed to a fake site.

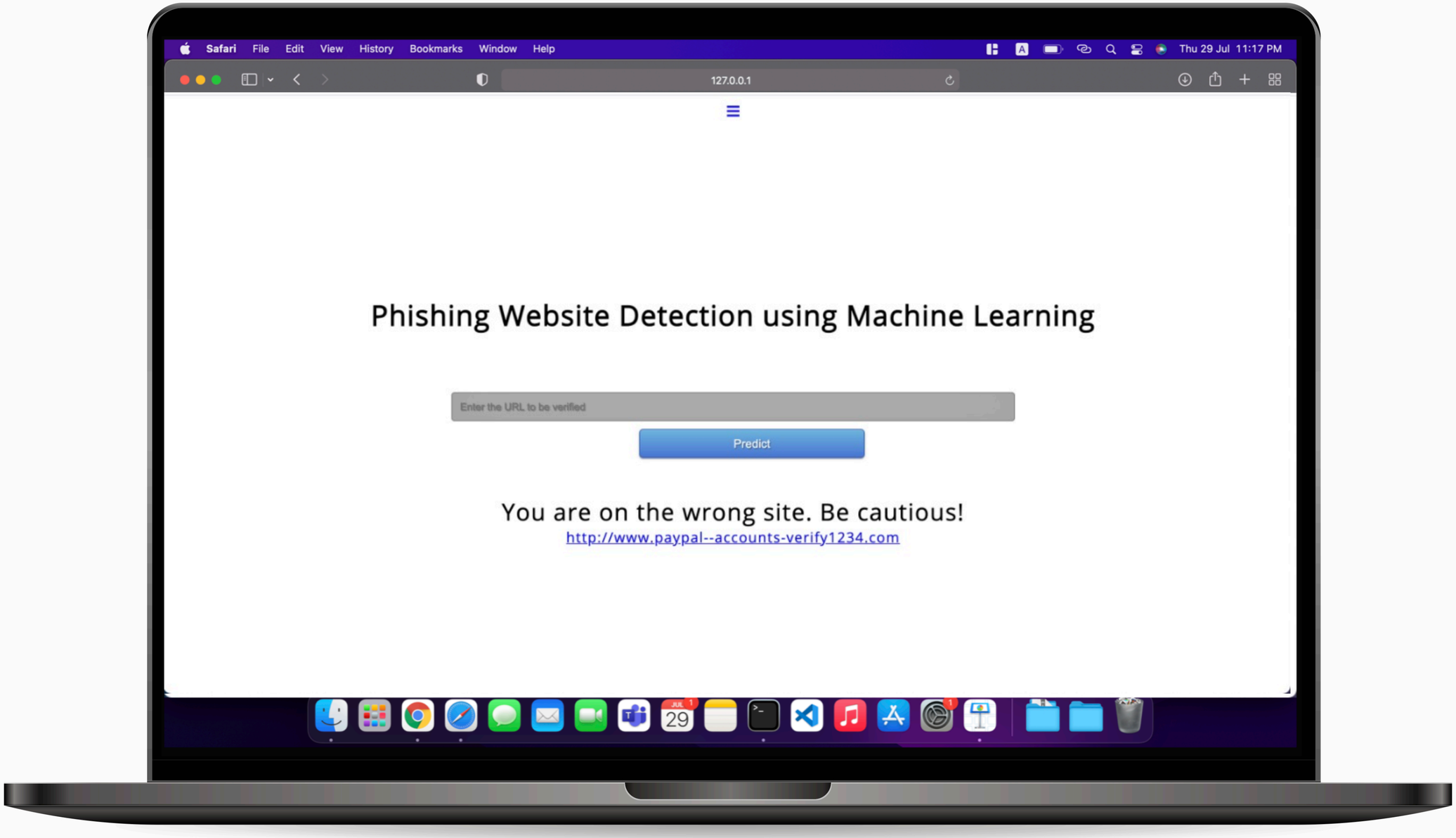
BROWSE SECURELY WITH HTTPS

You should always, where possible, use a secure website (indicated by https:// and a security "lock" icon in the browser's address bar) to browse, and especially when submitting sensitive information online, such as credit card details.


Safari File Edit View History Bookmarks Window Help
Thu 29 Jul 11:21 PM
127.0.0.1
Phishing Website Detection using Machine Learning
https://www.google.com
Predict


Safari File Edit View History Bookmarks Window Help
Thu 29 Jul 11:17 PM
127.0.0.1
Phishing Website Detection using Machine Learning
Enter the URL to be verified
Gmail
accounts.google.com/b/0/AddMailService
Predict
Your are safe!! This is a Legitimate Website.
https://www.google.com


Safari File Edit View History Bookmarks Window Help
Thu 29 Jul 11:30 PM
127.0.0.1
Phishing Website Detection using Machine Learning
http://www.paypal--accounts-verify1234.com
Predict


Safari File Edit View History Bookmarks Window Help
Thu 29 Jul 11:17 PM
127.0.0.1
Phishing Website Detection using Machine Learning
Enter the URL to be verified
Predict
You are on the wrong site. Be cautious!
http://www.paypal--accounts-verify1234.com

Future Work

- This application can be inserted into various applications regarding e-banking website.
- The UI of the web Application can be developed in variety of ways to look it more attractive
- In near future this module of prediction can be integrated with the module of automated processing system.

“So here, it can be concluded with confidence that the Artificial Neural Network (ANN) model is extremely efficient and gives a better result and fulfills all requirements of e-banking website. It predicts the nature of the e-banking website.”

Conclusion

SmartInternz + IBM

Thank You