KALINGA INSTITUTE OF INDUSTRIAL TECHNOLOGY (KIIT)

Deemed to be University U/S 3 of UGC Act, 1956

## KIIT School of Computer Science

## Course: Computer Networks

## Project Topic:

# Suspicious Email Detection


**Team Members:**

| | |
|---|---|
| Abhishek Mallick | 21051706 |
| Soumyabrata Samanta | 21051436 |
| Abhrajit Das | 21051026 |
| Deepraj Bera | 21051302 |

# Table of Content

## Introduction

Develop a comprehensive understanding of various email threats

## Data Collection

Identify and collect a suitable dataset comprising both legitimate and suspicious emails

## Feature Extraction

Extract relevant features from the data to facilitate the detection process.

## Model Training

A Logistic Regression model is chosen and trained on the transformed text features.

## Model Evaluation:

The accuracy of the model is evaluated on both the training and test datasets.

## Conclusion

the skills and knowledge needed to address the evolving challenges posed by email-based cyber threats

# 1.0 Introduction

- In today's interconnected world, email remains one of the primary communication channels.
- However, the rise of cyber threats, including phishing and malware distribution through emails, has necessitated the development of robust techniques for detecting suspicious emails.
- This project aims to explore the intricacies of email-based cyber threats and equip students with the skills to design and implement a Suspicious Email Detection system using concepts from computer networks.

# Understanding Email Threats

01

## Phishing

Phishing is a form of social engineering and scam where attackers deceive people into revealing sensitive information or installing malware such as ransomware.

02

## SPAM

The name comes from a Monty Python sketch in which the name of the canned pork product Spam is ubiquitous, unavoidable, and repetitive.

01

## Ransomware

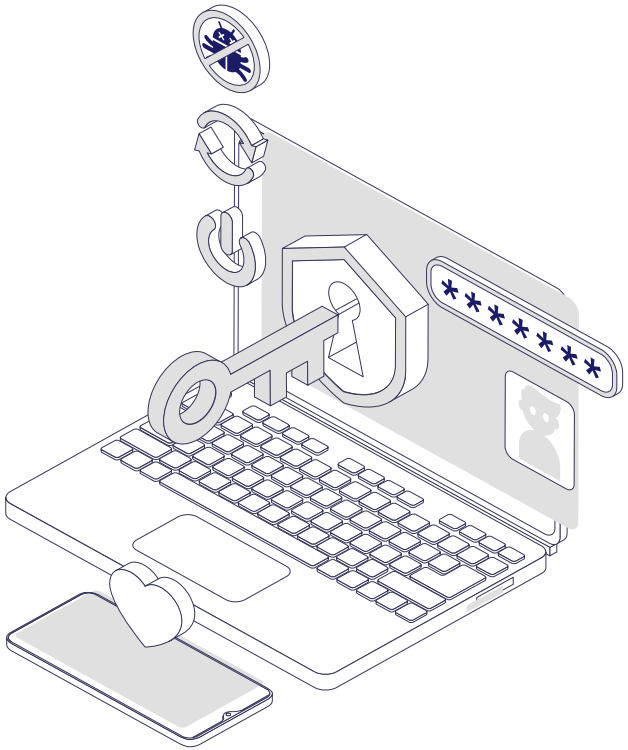
Ransomware is a type of cryptovirological malware that permanently block access to the victim's personal data unless a ransom is paid.

01

## Malware

Malware is any software intentionally designed to cause disruption to a computer, server, client, or computer network.

# 2.0 Dataset Acquisition:



**Goal**

Identify and collect a suitable dataset comprising both legitimate and suspicious emails for training and evaluation purposes. A diverse and representative dataset is crucial for effective model training.

**Dataset**

The importance of high-quality datasets cannot be overstated. The dataset serves as the bedrock upon which machine learning models are built, trained, and evaluated. Dataset link: Kaggle

**Code**

```
raw_mail_data = pd.read_csv('mail_data.csv')
```

# 3.0 Preprocessing and Feature Extraction

- Preprocess email data by addressing text, attachments, and headers.
- Extract relevant features from the data to facilitate the detection process.
- Proper preprocessing ensures the model's ability to discern patterns in the data.
- It replaces null values with an empty string.
- It then labels the data, assigning 'spam' emails a label of 0 and 'ham' (non-spam) emails a label of 1.

```
#Replace the null values with a null string
mail_data = raw_mail_data.where((pd.notnull(raw_mail_data)),‘’)
#Label spam mail as 0;  ham mail as 1;
mail_data.loc[mail_data['Category'] == 'spam', 'Category',] = 0
mail_data.loc[mail_data['Category'] == 'ham', 'Category',] = 1
```

# Data Splitting and Text Feature Extraction:

- The data is split into training and testing sets using the train_test_split function from scikit-learn.
- The text data is transformed into feature vectors using the TF-IDF vectorizer (TfidfVectorizer) from scikit-learn.
- This is a common technique in natural language processing to convert text data into numerical features.

```
#Separating the data as texts and label
X = mail_data['Message']
Y = mail_data['Category']

#Splitting the data into training data & test data
X_train, X_test, Y_train, Y_test = train_test_split(X, Y, test_size=0.2, random_state=3)

#Transform the text data to feature vectors that can be used as input to the Logistic
regression
feature_extraction = TfidfVectorizer(min_df = 1, stop_words='english', lowercase=True)

#Splited X has string values, those need to be fit & converted to integer
X_train_features = feature_extraction.fit_transform(X_train)
X_test_features = feature_extraction.transform(X_test)

#Convert Y_train and Y_test values as integers [convert object type to int]
Y_train = Y_train.astype('int')
Y_test = Y_test.astype('int')
```

# 4.0 Model Used

- **Logistic regression** is a statistical method for predicting a binary outcome. It uses mathematics to find the relationship between two data factors. The relationship is then used to predict the value of one of those factors based on the other.
- Logistic regression models can predict the probability of an outcome, event, or observation. The model delivers a binary or dichotomous outcome limited to two possible outcomes. For example, the model can predict the probability of passing an exam versus hours studying.

```
#Training the Model
model = LogisticRegression()

#Training the Logistic Regression model with the training data
model.fit(X_train_features, Y_train)
```

# 5.0 Model Evaluation

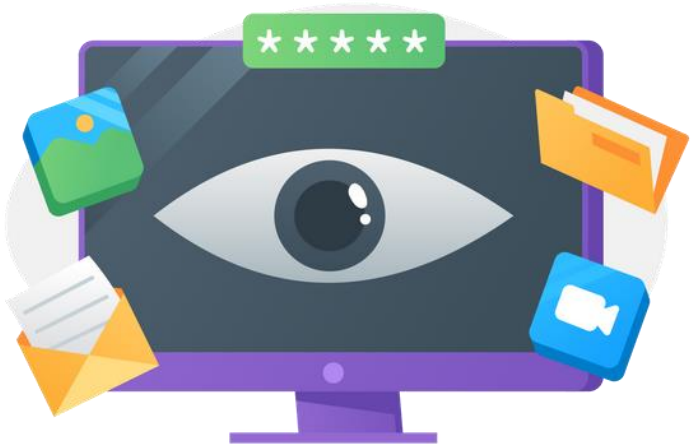
- A **Classification report** is used to measure the quality of **predictions** from a classification algorithm. How many predictions are True and how many are False. More specifically, True Positives, False Positives, True negatives and False Negatives are used to predict the metrics of a classification report.

Classification Report:

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| 0 | 1.00 | 0.75 | 0.86 | 155 |
| 1 | 0.96 | 1.00 | 0.98 | 960 |
| accuracy | | | 0.97 | 1115 |
| macro avg | 0.98 | 0.88 | 0.92 | 1115 |
| weighted avg | 0.97 | 0.97 | 0.96 | 1115 |

# 6.0 Conclusion

- By achieving these objectives, the project aims to equip individuals with the skills and knowledge needed to address the evolving challenges posed by email-based cyber threats.
- The combination of understanding threats, acquiring and processing data, selecting appropriate algorithms, designing effective models, and rigorous evaluation contributes to the development of a robust and reliable email threat detection system.



# References

https://scikit-learn.org/stable/index.html
https://pandas.pydata.org/
https://flask.palletsprojects.com/en/3.0.x/
https://www.ibm.com/topics/logistic-regression