

# Exploring the impact of AI on Social Engineering attacks

## AI for Cybersecurity Project

---

### Team Members:

Aayushi Sinha - 21BCS120, Mourya Kakarapu - 21BCS049, Udayini Vedantham - 21BCS130, Nithish Chouti - 21BCS074

---

### What is a Social Engineering attack?

A social engineering attack is one of the forms of cybersecurity attacks, where attackers use psychological tactics to bring the reluctant others to disclose confidential information, spread their login details, allow them to access their devices or otherwise compromise their digital security.<sup>[1]</sup>

An attacker might use social skills to gather or compromise information about an organization or its systems by appearing trustworthy and legitimate. They might pose as a new employee, technician, or researcher and may even present credentials to establish credibility. Through seemingly harmless questions, the attacker can gradually collect enough details to gain unauthorized access to the organization's network. If a single source doesn't provide enough information, the attacker may approach other individuals in the organization, using details from previous interactions to strengthen their credibility and gain further access.<sup>[2]</sup>

### Psychological tactics in Social Engineering

The social engineers of the world are quite dependent on triggering the brain in a way that prompts people to act without thinking too much about it. Here's generally how they do it:<sup>[3][4]</sup>

- **Creating urgency:** Attackers may build a fake sense of urgency that he/she has to act immediately. For example, it may say something like the following: "Please respond as your account will be locked within 10 minutes if no reply is received." The idea is to pressure the individual into pressing action before having had a chance to verify the message.
- **Impersonation:** Attackers create an illusion of trustworthiness by masquerading as someone you know: a bank official, government representative, or tech support, to make their demands seem legitimate. This is because requests that seem familiar or from a trusted source get more responses.
- **Exploiting curiosity:** Many social engineering attacks exploit curiosity to motivate someone to take action, such as with a subject line that says, "You have won a prize! or "Read this secret document for your mystery prize."
- **Intimidation and fear:** Threats of possibly punitive actions to be taken against their targets if the latter do not heed their demands, such as cutting off access or prosecuting, usually makes them divulge confidential information.
- **Building rapport:** This is a scenario where the attackers pretend to be an acquaintance. This builds trust hence the likelihood of the victim divulging information increases.

## Common Social engineering attack types

The most common types of social engineering attacks are:[5]


1. **Phishing:** This is the most common type of social engineering where the attacker sends fraudulent emails that mimic legitimate communications from trusted entities to trick recipients into divulging sensitive information, such as login credentials or credit card numbers.
2. **Spear phishing:** This is an advanced form of phishing. It targets precise individuals or organizations. Email is more personalized, gathered from social media or perhaps other places to make such emails seem more authentic or credible.
3. **Vishing and Smishing:** Vishing is voice phishing, or SMS phishing, which calls or texts people to coax them into revealing personal and financial information.
4. **Whaling:** Very similar to spear phishing but involves targeting only high profile individuals in organizations.
5. **Pretexting:** In this the attackers create an imaginary scenario to attain access to sensitive information. For example posing as an IT technician needing login credentials to perform maintenance.
6. **Baiting:** This involves offering something enticing or lucrative to tempt targets into revealing information or downloading malware. For example, an attacker might claim to show an exclusive link to the latest movie which would actually be a compromised website.

## Role of AI in enhancing Social Engineering

Social engineering attacks play on the victim's psychology, exploiting them through manipulation. Traditionally, these attacks were carried out through manual means, mainly mass emails or calls. With the advancement of AI and NLP, these attacks are being automated and personalized, posing a greater danger than ever.


The increasing use of Generative AI models has allowed attackers to create personalized and genuine phishing messages, enabling them to impersonate legitimate entities, manipulate people through deepfake and extort sensitive information. “These models, [like] ChatGPT, FraudGPT, and WormGPT, have augmented existing threats and ushered in new dimensions of risk”<sup>[6]</sup>. Thus it is important to understand and analyze how these models are used in such attacks, in order to create strategic countermeasures and AI-powered security systems.

In a research published by Harvard Business Review<sup>[7]</sup>, they have shown how 60% participants of the research fell for AI-generated automated phishing attacks. Further, the researchers reported that their new findings show the entire phishing process can be automated using large language models (LLMs), reducing the costs of phishing attacks by more than “95% while achieving equal or greater success



संस्कृतम्  
विद्यया ऽमृतम्  
भारतम्  
MINISTRY OF  
HEALTH AND  
FAMILY WELFARE  
संस्कारे ऽमृतम्

# Cyber Security Awareness



## Beware of FraudGPT Scam

FraudGPT, an AI-powered Chatbot is used by Cyber criminals to craft fraudulent content for cyber frauds and crimes.


### Modus Operandi

- FraudGPT can generate authentic-looking phishing emails, text messages, or websites that trick users to reveal sensitive information, such as login credentials, financial details, or personal data.
- It can create deceptive messages to trick users to click on malicious links/attachments leading to malware infections.
- It can imitate human conversation with users to share sensitive information or to perform harmful actions.
- It can help hackers create fraudulent documents, invoices, or payment requests for financial scams.


### Safety Tips

- Avoid clicking on links/ attachments from unknown sources.
- Always verify the authenticity of calls, emails or messages, especially those asking for sensitive information or financial transactions.
- Contact the organization directly through their official channels to validate such requests.
- Regularly update security software, install patches, and use genuine antivirus programs to protect against potential threats.


For more safety tips visit: <https://www.cert-in.org.in> and <https://www.csk.gov.in>




Department of Information Technology  
Ministry of Electronics and Information Technology




Ministry of Information and Public Relations




Ministry of Information and Public Relations




Ministry of Information and Public Relations




Ministry of Information and Public Relations



Ministry of Information and Public Relations



Ministry of Information and Public Relations



Ministry of Information and Public Relations

rates”. This signifies the looming threat possessed by the use of AI in phishing attacks, which is predicted to be increasing and becoming even stronger with time.

*“Scammers see AI tech as a gold mine for phishing schemes”* - quoted in an article by Microsoft<sup>[8]</sup>.

### Traditional vs AI-powered phishing attacks

Traditional phishing attacks often contain many spelling mistakes, grammar issues and awkward phrasings, making them easier to categorize as illegitimate. Using AI, these errors are highly minimized to the extent that the messages appear legitimate and accurate enough to win the user's trust. Large language models (LLMs) can also pull in current information from news and company websites, making phishing emails more relevant and urgent. AI chatbots help create and spread targeted phishing attacks, like business email compromise and whaling, much faster than human attackers could which greatly increases the scale and impact of these attacks.

In an experiment conducted by Singapore’s Government Technology Agency, the security team sent spear phishing emails to their internal users in which some were written by humans, while others were generated by OpenAI’s ChatGPT-3 Model. “More people clicked the links in the AI-generated phishing emails than in the human-written ones -- by a significant margin.”<sup>[9]</sup>

### Let’s take an example

Traditional phishing emails:

#### Attention! Your PayPal account will close soon!

Dear Member,

We have faced some problems with your account Please update the account .If you do not update will be Closed.

To Update your account, just confirm your informations.(It only takes a minute.)

It's easy:

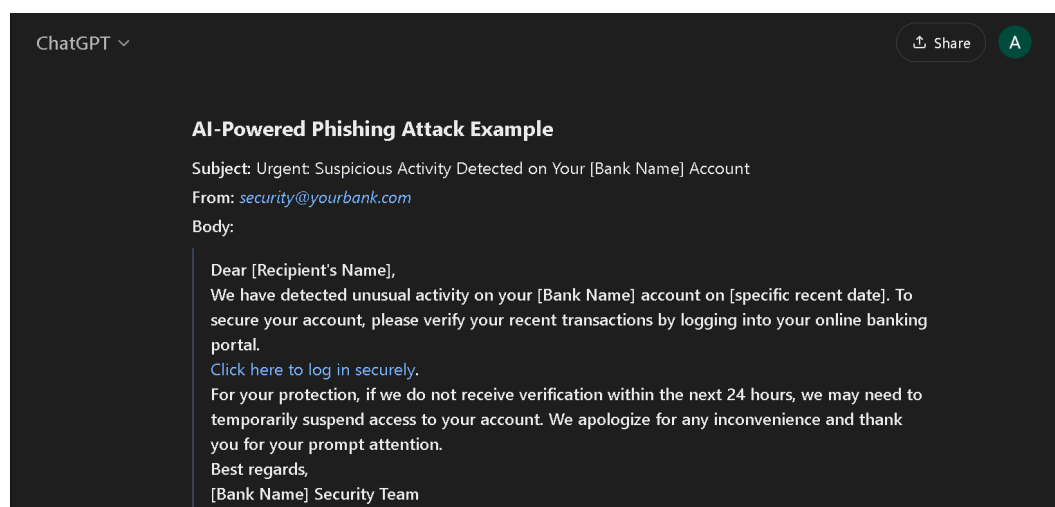
1. Click the link below to open a secure browser window.
2. Confirm that you're the owner of the account, and then follow the instructions.

**Relog in your account now**

This is a common spam mail that we have all received at least once. To anyone alert and cautious, it is quite easy to identify this mail as a phishing email.

The mail has many spelling and grammatical errors like “Relog” instead of “LogIn”. It creates a sense of urgency for the user, including a threat (though, poorly phrased). The link probably directs the user to malicious sites that could install harmful malwares into their system or steal their data.

Now let’s look at an example of phishing email crafted using OpenAI’s ChatGPT-3, prompted by us:



In contrast to the previous mail, this mail is much harder to categorize as phishing or not. It appears to be legitimate or closely crafted to actual mails sent by banks. With phrasings such as “Click here to log in securely.” and “ We apologize for any inconvenience”, it does well in manipulating a regular user. The perpetrator behind the attack would only need to do a basic research in personalizing the message further for the user, like adding the bank name or the user’s name.

Even more concerning is the selling of subscription-based “FraudGPT” and “WormGPT” models in the Dark Web, as found by researchers<sup>[10]</sup>, which is aimed at aiding hackers in their malicious activities. This posits the urgent need for extreme countermeasures and AI-powered cybersecurity defenses.

It raises a critical question: how can we design a model that accurately detects phishing emails and protects users from these evolving threats? The first step is identifying and analyzing behavioral patterns used for such purposes.

## Identifying behavioral patterns used in social engineering attacks

According to the Cybersecurity and Infrastructure Security Agency (CISA), here are 6 common indicators of social engineering attacks and ways to recognize them:

### 1) Suspicious sender’s address

Cybercriminals will often imitate the address of a legitimate business when sending you an email or a message. The sender's address may closely resemble one from a reputable company, but with some characters altered or omitted.

### 2) Generic greetings and signature

Usually, a generic greeting like “Dear Valued Customer” or “Sir/Ma’am”, combined with a lack of contact information in the signature block, strongly indicate phishing. That’s because a legitimate email from a trusted organization will normally provide their contact information and address you by name and/or honorific.

### 3) Spoofed hyperlinks and websites

Spoofed links can be easily identified if you hover your cursor over any of the links in the body of the email. If the links do not match the text that appears when you hover over them, that's an indication that the link may be spoofed. Malicious websites can also look identical to a legitimate site, but when you check the URL, it uses a variation in site's spelling or a different domain (i.e., a government site with a .net domain instead of .gov). Moreover, cybercriminals may shorten their URLs to hide the true destination of the link.

#### **4) Secondary Destinations**

Some phishing attacks involve directing the victim to a legitimate document hosting site, or attaching a non-malicious document to the message. In other words, bad actors can insert a message with a link within the harmless document. This will direct the victims into the malicious site, where the actor hosts infected files or a credential skimming scam.

#### **5) Spelling and layout**

This is one of the most obvious indicators of a possible phishing attack — a message with poor grammar and sentence structure, misspellings, and inconsistent formatting. That's because reputable institutions almost always have personnel dedicated to producing, verifying, and proofreading their customer correspondence.

#### **6) Suspicious attachments**

Unsolicited emails requesting the user to download and/or open an attachment commonly indicate a malware attack. Too often, a cybercriminal uses a false sense of urgency or importance to persuade the user to download/open the attachment without examining and confirming first. E.g., a bad actor may pretend to be an executive and say "I need this document printed and on my desk in 10 minutes," or something to that effect.

#### **7) An Offer Too Good to Be True**

Cybercriminals sometimes use blackmail or coercion to extract information, especially during the information gathering stage of a social engineering attack.

### **How can natural language processing be used to analyze social engineering tactics**

Natural language processing (NLP) is a subset of artificial intelligence (AI) that deals with the interaction between computers and humans in natural language. In the context of social engineering, NLP can be used to analyze tactics by extracting insights from large amounts of text data, such as emails, chat logs, and social media posts. This can help identify patterns and trends in social engineering attacks, enabling organizations to develop more effective defense strategies.

#### **Here are some ways NLP can be used to analyze social engineering tactics:**

- 1) Text classification: NLP can be used to classify social engineering tactics into different categories, such as phishing, spear phishing, or business email compromise (BEC).
- 2) Named entity recognition (NER): NLP can be used to identify and extract specific entities, such as names, locations, and organizations, from text data.

- 3) Part-of-speech tagging: NLP can be used to identify the part of speech (such as noun, verb, adjective, etc.) of each word in a sentence, which can help identify the intent behind the text.
- 4) Sentiment analysis: NLP can be used to analyze the sentiment of text data, such as whether it is positive, negative, or neutral.
- 5) Topic modeling: NLP can be used to identify underlying topics in large amounts of text data, such as common themes or ideas.

**A possible implementation of an NLP-based social engineering defense system could look like this:**

- 1) Data Collection: Gather datasets of emails and messages labeled for phishing, spam, and legitimate communication. Add annotations for behavioral cues like urgency or authority.
- 2) Feature Extraction and Preprocessing: Preprocess text using tokenization, lemmatization, and feature extraction techniques. Sentiment scores, entity names, and specific keywords or phrases are extracted as key features.
- 3) Model Training: Train models like BERT for contextual analysis, using labeled data for phishing and legitimate communication. Fine-tune on custom datasets to increase accuracy for specific organizations.
- 4) Deployment and Integration: Integrate NLP models into communication channels (e.g., email servers) to perform real-time analysis, flagging suspicious messages.

**Some of the benefits of using NLP to analyze social engineering tactics include:**

- 1) Improved incident response:** By analyzing social engineering tactics, organizations can develop more effective incident response plans and procedures.
- 2) Enhanced threat intelligence:** NLP can provide insights into the tactics, techniques, and procedures (TTPs) used by social engineers, enabling organizations to stay ahead of emerging threats.
- 3) Increased efficiency:** NLP can automate the process of analyzing large amounts of text data, freeing up human analysts to focus on higher-level tasks.

**Some of the challenges of using NLP to analyze social engineering tactics include:**

- 1) Data quality:** NLP requires high-quality data to produce accurate results, but social engineering data can be noisy and difficult to collect.
- 2) Contextual understanding:** NLP can struggle to understand the context of social engineering tactics, which can lead to false positives or false negatives.
- 3) Scalability:** NLP can be computationally intensive, making it challenging to analyze large amounts of data in real-time.

## **AI/ML Based Solutions to detect and mitigate Phishing Attacks**

- 1) Phishing Text Classification Detection:** Analyzing the content of an email, sms and text messages via NLP and algorithms like Naive Bayes or Logistic Regression in order to classify whether it's a phishing activity or not.
- 2) Phishing URL Detection:** We can consider characteristics of the website like length of URL, usage of special characters, etc to determine whether it is phishing or not.

**3) Real-time Suspicious Activity Monitoring:** We can monitor and detect unusual patterns like spikes in login attempts or irregular user behaviors to identify possible phishing attacks. For this purpose, we can use time series analysis and anomaly detection.

**4) Human-Machine Interaction Monitoring:** Monitoring interaction between human and automated systems by sentiment and intent analysis using NLP, behavior modeling and monitoring chatbot response logs.

## **Solutions Implemented by us:**

### **1) Phishing Text Classification Model**

Dataset: <https://www.kaggle.com/datasets/uciml/sms-spam-collection-dataset>

Colab Link:

[https://colab.research.google.com/drive/1VhGvf6YjszBgbNw1Y7ka3JLbPMphdY4\\_?usp=sharing](https://colab.research.google.com/drive/1VhGvf6YjszBgbNw1Y7ka3JLbPMphdY4_?usp=sharing)

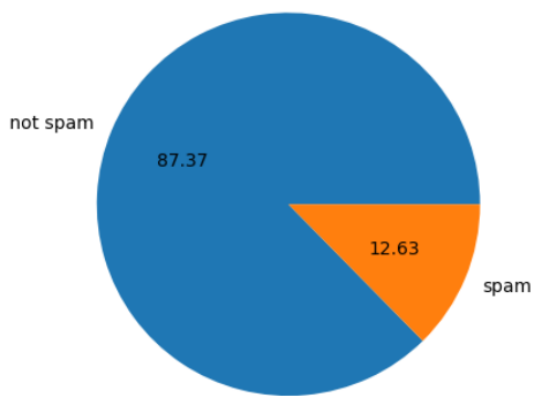
## **Email and SMS Phishing in Relation to Text Classification Using NLP and AI/ML**

Phishing is a deceptive cybercrime where attackers attempt to manipulate individuals into revealing sensitive information, such as passwords or financial details, through fraudulent emails or SMS messages. These malicious messages often employ urgent language, suspicious links, or enticing offers to lure unsuspecting users.

In the fight against phishing, Natural Language Processing (NLP) combined with Artificial Intelligence (AI) and Machine Learning (ML) plays a critical role. These technologies enable automated classification of text messages by analyzing their linguistic patterns and identifying traits indicative of spam. By leveraging NLP for feature extraction and ML for prediction, we can build systems capable of distinguishing between legitimate (ham) and phishing (spam) messages with remarkable precision.

### **Dataset Overview**

The dataset used for this study, the SMS Spam Collection, is a benchmark dataset frequently employed in spam detection research. It comprises 5,574 SMS messages in English, each labeled as either ham (legitimate) or spam. This dataset provides a reliable basis for training and evaluating machine learning models. With a diverse collection of real-world messages, it captures the linguistic nuances of both spam and legitimate communication, making it ideal for robust text classification tasks.



message_type		message
2	1	Free entry in 2 a wkly comp to win FA Cup fina...
5	1	FreeMsg Hey there darling it's been 3 week's n...
8	1	WINNER!! As a valued network customer you have...
9	1	Had your mobile 11 months or more? U R entitle...
11	1	SIX chances to win CASH! From 100 to 20,000 po...
...	...	...
5537	1	Want explicit SEX in 30 secs? Ring 02073162414...
5540	1	ASKED 3MOBILE IF 0870 CHATLINES INCLU IN FREE ...
5547	1	Had your contract mobile 11 Mnths? Latest Moto...
5566	1	REMINDER FROM O2: To get 2.50 pounds free call...
5567	1	This is the 2nd time we have tried 2 contact u...

Figure Depicting the Overview of the Dataset

## Features Extracted from the Dataset

Effective text classification depends on the identification and representation of meaningful features. The following features were considered for this project:

- 1. Message Text:** The raw text of each SMS serves as the foundational input for analysis.
- 2. Word-Level Metrics:** Frequency distributions of words, including n-grams (bigrams and trigrams), to capture context.
- 3. Message Length:** A numerical feature reflecting the total character or word count in the SMS.
- 4. TF-IDF Scores:** Term Frequency-Inverse Document Frequency (TF-IDF) measures the importance of words relative to their occurrence in the dataset.
- 5. Special Patterns:** Detection of specific patterns such as URLs, phone numbers, and promotional phrases that are common in spam messages.

These features provide a comprehensive understanding of each message, facilitating the development of a robust classification model.

## NLP Techniques Employed

To transform raw text into a format suitable for machine learning models, several NLP techniques were utilized:

- 1. Tokenization:** Each SMS was split into smaller units (words or tokens) to enable feature extraction.
- 2. Stopword Removal:** Commonly used words that do not carry significant meaning (e.g., "and," "the") were excluded to reduce noise.
- 3. Stemming and Lemmatization:** Words were reduced to their base or root forms (e.g., "running" to "run") to improve consistency in analysis.
- 4. Vectorization:** The text was converted into numerical representations using methods such as:
  - **Bag of Words (BoW):** Encodes the frequency of each word in a message.



- **TF-IDF**: Balances the frequency of words with their relevance across the dataset.

These techniques allowed the conversion of unstructured textual data into structured formats, ensuring compatibility with machine learning algorithms.

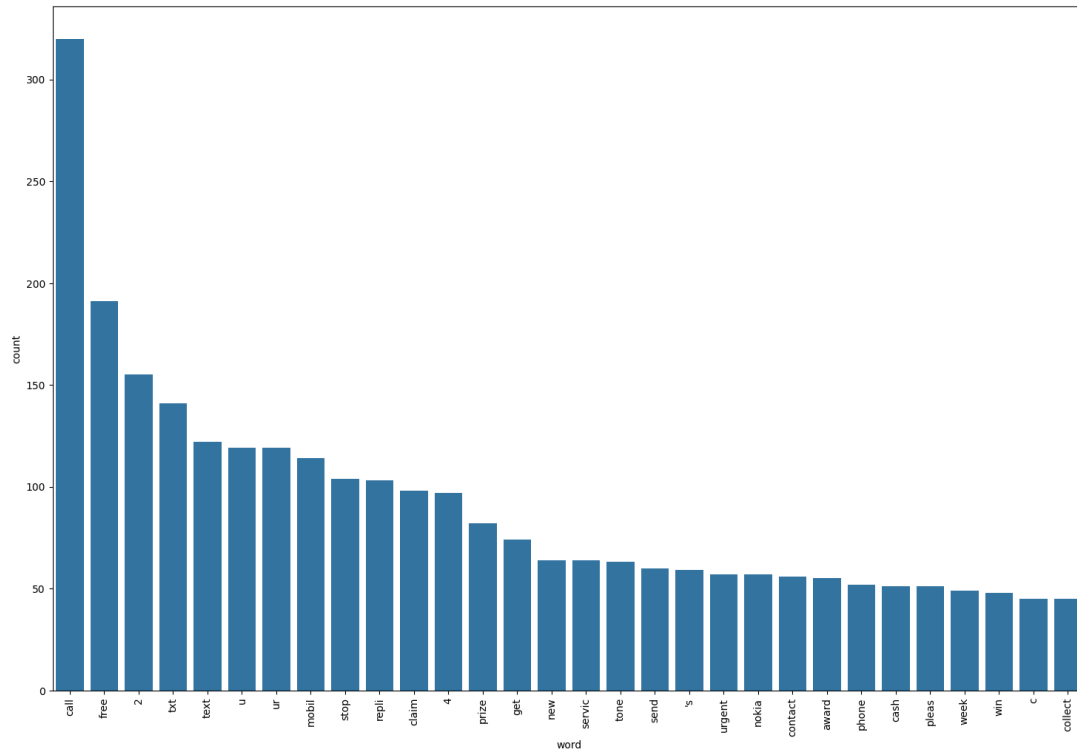


Figure Depicting the Most Frequent words Occurring in the Spam Messages

## Preprocessing Pipeline

The preprocessing phase is a critical step in preparing the data for model training. The following steps were taken to clean and organize the dataset:

- 1. Text Cleaning:** Removed special characters, punctuation, and unnecessary whitespace from the messages.
- 2. Case Normalization:** Converted all text to lowercase to ensure uniformity.
- 3. Stopword Filtering:** Eliminated common, non-informative words to emphasize meaningful content.
- 4. Numerical Feature Transformation:** Applied vectorization techniques like TF-IDF to encode the text as numerical arrays.
- 5. Data Splitting:** Divided the dataset into training (80%) and testing (20%) sets to assess model performance.

This pipeline ensured that the input data was both accurate and efficient for training machine learning models.

## Naive Bayes Model

For this project, two variants of the Naive Bayes classifier were implemented: Gaussian Naive Bayes (GNB) and Multinomial Naive Bayes (MNB). The Naive Bayes algorithm, grounded in Bayes' Theorem, is a probabilistic model well-suited for text classification tasks such as spam detection.

**Gaussian Naive Bayes (GNB):** Assumes that the features follow a Gaussian (normal) distribution. While this assumption works well for continuous data, it has limitations for text-based datasets, where word frequencies are typically non-Gaussian.

**Multinomial Naive Bayes (MNB):** Designed specifically for text data, it models word counts or term frequencies using a multinomial distribution, making it highly effective for spam detection tasks. Both models were evaluated on the SMS Spam Collection dataset, offering valuable insights into their strengths and limitations.

## Results and Model Performance

The performance of both Gaussian Naive Bayes and Multinomial Naive Bayes classifiers was evaluated using metrics such as accuracy, precision, and confusion matrices:

Gaussian Naive Bayes (GNB)

Accuracy: 87.81%

Precision: 52.68%

Confusion Matrix:

[ 790 106 ]

[ 20 118 ]

The GNB model struggled with text classification due to its assumption of Gaussian-distributed features, leading to lower precision and a higher false-positive rate.

Multinomial Naive Bayes (MNB)

Accuracy: 97.49%

Precision: 100%

Confusion Matrix:

[ 896 0 ]

[ 26 112 ]

The MNB model outperformed GNB, achieving near-perfect accuracy and precision. It successfully identified spam messages with zero false positives, though a small number of spam messages were misclassified as legitimate.

## Comparative Insights

The results clearly demonstrate the superiority of the Multinomial Naive Bayes model for text-based spam detection. Its design, tailored for discrete features like word frequencies, aligns closely with the nature of the dataset, enabling higher accuracy and precision compared to Gaussian Naive Bayes.

### Sample Test Cases to show Model Efficiency:

```
input_sms = """Subject: Urgent: Your Account Needs Immediate Attention

Dear User,

We have noticed some unusual activity on your account and need to verify your identity to prevent any potential security threats. To ensure the safety of your account
[Verify your account now]

Failure to complete this verification within the next 24 hours will result in your account being temporarily suspended for your protection.

If you did not initiate this activity, please contact our support team immediately.

Thank you for your cooperation.

Sincerely,
The Support Team"""

transformed_sms = transform_text(input_sms)
# 2. vectorize
vector_input = tfidf.transform([transformed_sms])
# 3. predict
result = model.predict(vector_input)[0]
# 4. Display
if result == 1:
    print("Spam")
else:
    print("Not Spam")
```

Spam

Figure showing an example of Spam Email.

```
input_sms = """Dear Sir,

I hope this message finds you well. I am Udayini Vedantham, 4th year CSE and am reaching out to request a letter of recommendation from you as I prepare my applicati
I have the privilege of working under your guidance for the course on Information Retrieval and its project Information retrieval from Text and Images along with the
I would be happy to provide any additional information or meet at your convenience to discuss this further. Additionally, I can share details about the universities
Thank you very much for your consideration sir

Udayini Vedantham,
21BCS130"""

transformed_sms = transform_text(input_sms)
# 2. vectorize
vector_input = tfidf.transform([transformed_sms])
# 3. predict
result = model.predict(vector_input)[0]
# 4. Display
if result == 1:
    print("Spam")
else:
    print("Not Spam")
```

Not Spam

Figure showing an example of Ham Email.

## 2) Phishing URL Detection

Dataset: <https://www.kaggle.com/datasets/shashwatwork/web-page-phishing-detection-dataset/data>

Colab Link:

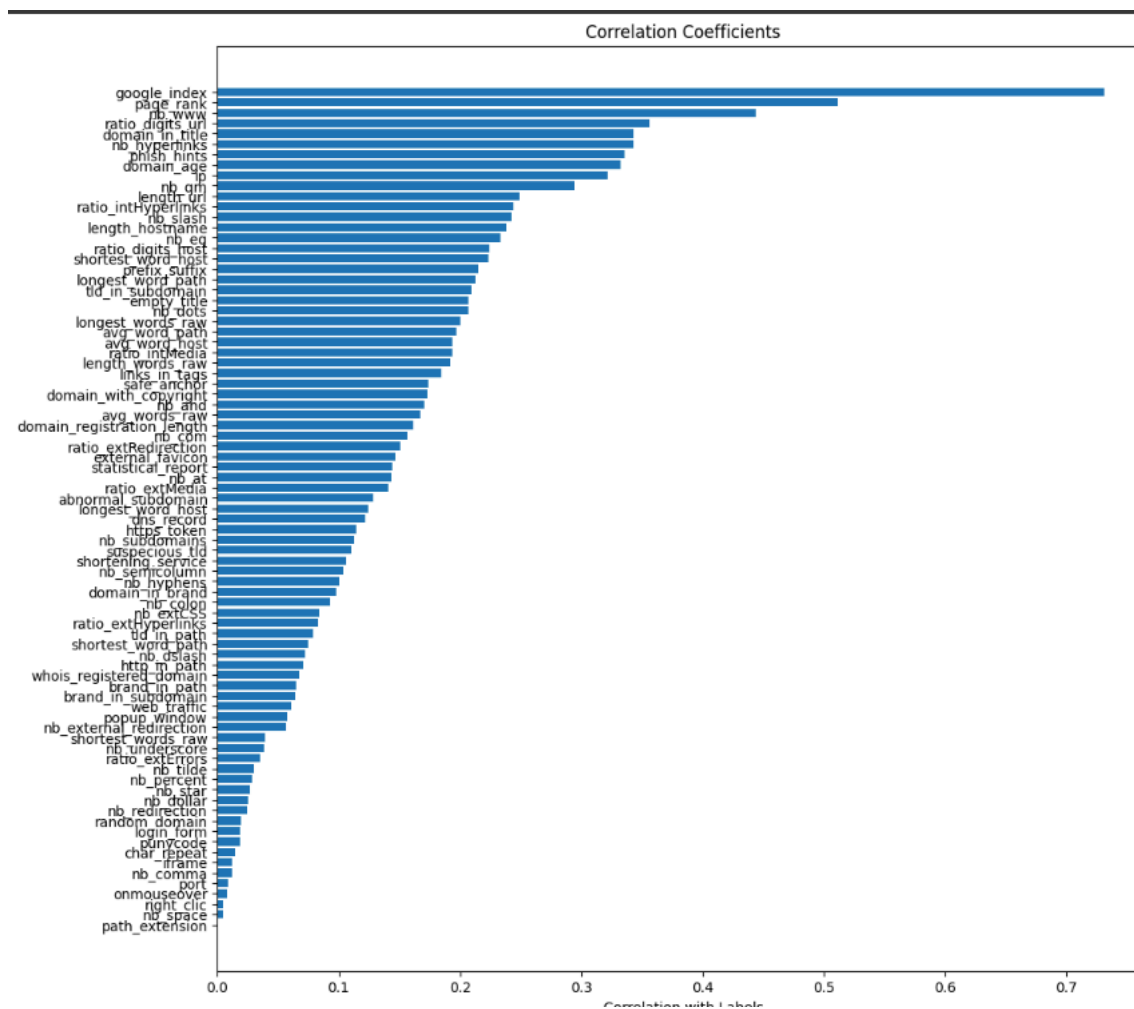
[https://colab.research.google.com/drive/1yNcH74s2ndhJVeEurXv3ZRRhcYfpzl8y#scrollTo=avgnnNQkhJ\\_v](https://colab.research.google.com/drive/1yNcH74s2ndhJVeEurXv3ZRRhcYfpzl8y#scrollTo=avgnnNQkhJ_v)

Phishing mails often contain links to malicious websites, hence detecting these links is crucial along with text classification to create a more robust security system.

Our dataset contains 5715 Legitimate websites and 5715 Phishing websites data along with 80 features extracted from these websites such as Google Index, Page Rank, Number of hyperlinks, Domain Age, Length of URL etc.

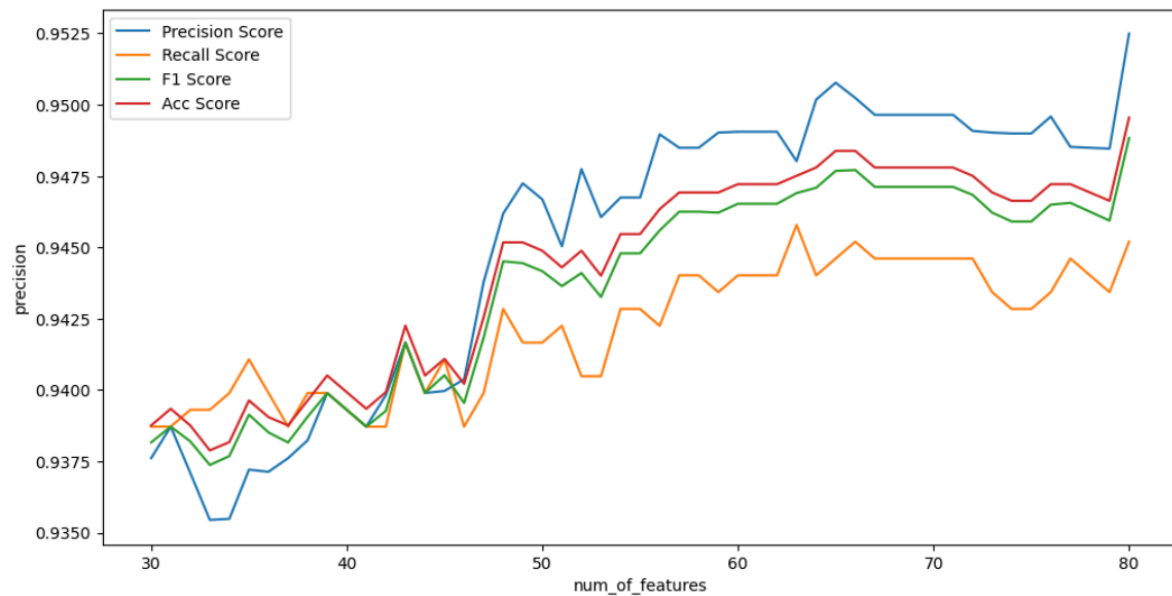
We created a correlation matrix to understand the relationship between the features and the Label, in order to understand which features are more relevant and contribute better to the classification model.

We ranked these features based on absolute correlation coefficients, to get the top features list.

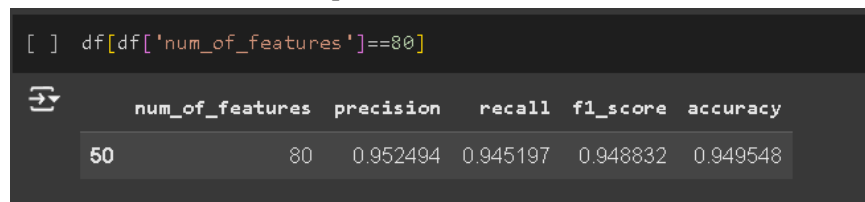


Further, we applied two algorithms - Logistic Regression and Random Forest Classification Models, to train the model and select the best performing one. For each of these models, we began by training with the top 30 features to top 80 features, to determine the number of top features required to obtain the best results.

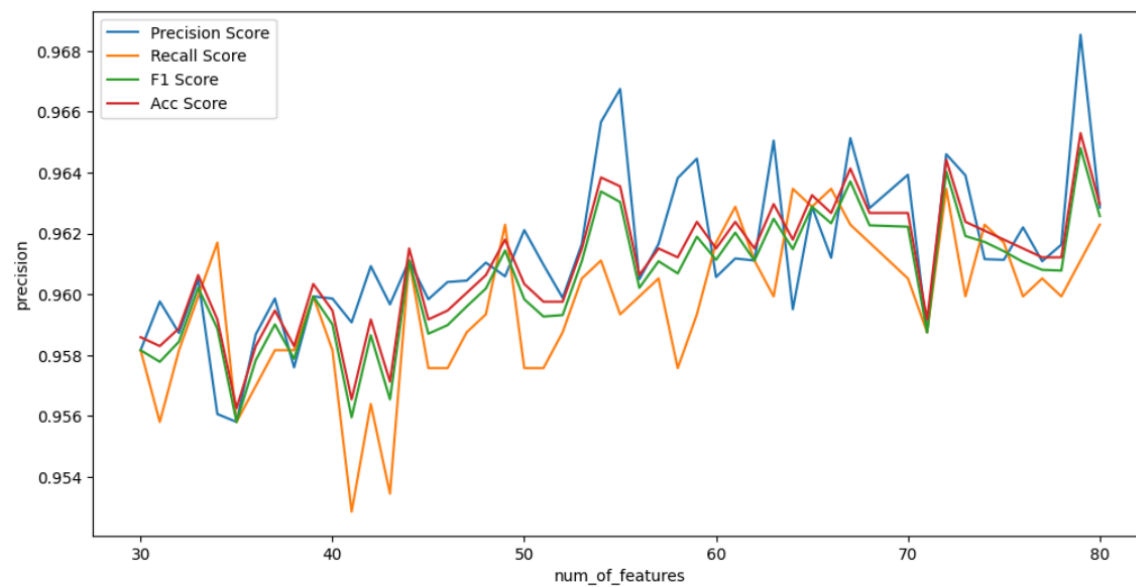
## Logistic Regression:



Best model is obtained at top 80 features:



## Random Forest Classification:



Best model is obtained at Top 79 Features:

```
[ ] df_rf[df_rf['num_of_features']==79]
```



	num_of_features	precision	recall	f1_score	accuracy
49	79	0.968527	0.961108	0.964803	0.965296

Random Forest outperforms Logistic Regression, so we trained it on the top 79 features. Finally we conducted inference on our validation set by taking any random legitimate and phishing website and predicting it using our model:

```
legitimate_prediction = rf.predict(legitimate_row_scaled)
phishing_prediction = rf.predict(phishing_row_scaled)

print("Legitimate Website Prediction: ", legitimate_prediction[0], " | Actual Label: 0")
print("Phishing Website Prediction: ", phishing_prediction[0], " | Actual Label: 1")
```



```
Legitimate Website Prediction:  0 | Actual Label: 0
Phishing Website Prediction:  1 | Actual Label: 1
```

We obtained accurate results for the validation set.

## References

- [1]<https://www.crowdstrike.com/en-us/cybersecurity-101/social-engineering/types-of-social-engineering-attacks/>
- [2]<https://www.cisa.gov/news-events/news/avoiding-social-engineering-and-phishing-attacks>
- [3]<https://abnormalsecurity.com/glossary/social-engineering>
- [4]<https://votiro.com/blog/5-psychological-tricks-hackers-use-to-make-users-download-files/>
- [5]<https://www.techbrain.com.au/social-engineering-attack-psychology/>
- [6] <https://arxiv.org/pdf/2310.05595>
- [7] <https://hbr.org/2024/05/ai-will-increase-the-quantity-and-quality-of-phishing-scams>
- [8]<https://www.microsoft.com/en-us/microsoft-365-life-hacks/privacy-and-safety/how-ai-changing-phishing-scams>
- [9]<https://www.techtarget.com/searchsecurity/tip/Generative-AI-is-making-phishing-attacks-more-dangerous>
- [10] <https://secureops.com/blog/ai-attacks-fraudgpt/>
- [11]<https://massedcompute.com/faq-answers/?question=How%20can%20natural%20language%20processing%20be%20used%20to%20analyze%20social%20engineering%20tactics?>
- [12]<https://www.safeguardcyber.com/identify-prevent-social-engineering-attacks#Common-Indicators-of-Social-Engineering-Attacks>