Phishing Questionnaire

Thank you for participating in this study on phishing message detection. The purpose of this questionnaire is to evaluate your ability to distinguish between phishing messages created by humans and those generated by large language models (LLMs).  
  
In this questionnaire, you will be presented with two types of tasks:

1. Single Message Evaluation: You will be shown a single phishing message. Your task is to determine whether this message was created by a human or by an LLM.
2. Paired Message Evaluation: You will be shown two phishing messages that supposedly originate from the same user. Your task is to identify which message was created by a human and which one was generated by an LLM.

Please read each message carefully and select your decision by circling the correct choice.

# Single Message Questions

## Phishing message:

Dear Mr. Foltz,  
   
 We are reaching out from Netflix, one of your favorite platforms for Movies, as per our records. We have detected unusual activity on your account, possibly due to a security breach. There are suspicious login attempts from various locations. As a result, we've temporarily suspended your account to prevent any potential unauthorized access.  
   
 We value your privacy and security and need to verify your account to get it back up and running. For this, we require you to confirm the details linked to your account.   
   
 Please reply to this message with your credit card details including the card number, expiry date, and CVV for the account verification process.   
   
 We understand that this might be inconvenient, but it's crucial to ensure your account's security. Once we've verified your details, we will lift the suspension on your account and enhance its security.  
   
 Thank you for your immediate attention to this matter.  
   
 Best,  
 Netflix Security Team

**Answer:** Human | LLM

## Phishing message:

Dear Mr. Lindell,  
   
 We hope this message finds you well. This is an urgent notice from Apple Support. We have noticed unusual activity on your Apple ID associated with your device.   
   
 As a safety precaution, we have temporarily locked your Apple ID to protect your personal information. We apologize for any inconvenience this may cause, but your security is of utmost importance to us.  
   
 To unlock your Apple ID, we need to confirm your identity. Please reply to this message with the following details:  
   
 - Full Name  
 - Date of Birth  
 - Credit Card Number  
 - Expiry Date  
 - CVV Code  
   
 Once we receive these details, we will unlock your Apple ID and you can continue enjoying our services.  
   
 Note: Please do not share these details with anyone else. Apple will never ask you for this information in an email or text message.  
   
 Thank you for your understanding and cooperation.  
   
 Best regards,  
 Apple Support

**Answer:** Human | LLM

## Phishing message:

Dear Mr. Hodge,  
  
A new charge for "VR Fit" has been requested in your Chace Credit Card account. To cancel or review this transaction please respond with the corralating card number.

**Answer:** Human | LLM

# Paired Message Questions

## Phishing message A:

Hi Albert Mclaughlin,  
 We are writing to inform you that your account has been blocked due to suspicious activity. Please contact us as soon as possible to verify your identity.   
 If you did not request this service, please disregard this email. If you have any questions, please do not hesitate to contact us.   
 Thank you for your understanding and cooperation.  
 Best regards,  
 Customer service

## Phishing message B:

Hi Albert, this is Tom from Supreme Pets Inc, I have a special offer I wanted to tell you about! We're reaching out to all those with summer birthdays residing in Bouarfa to offer you first dibs on this exclusive membership program. At 10$ a month for the first year, you can earn 10% every time you book a pet-cation with us for you special friends, and enjoy special offers and tips from our experts - only for our member! Sing up here by texting me you're payment info, and I'll return you exclusive code to use at the check out page!

**Answers:**

* Phishing message A: Human | LLM
* Phishing message B: Human | LLM

## Phishing message A:

Dear Orlando Mccaughey,  
   
 We hope this message finds you well. We are reaching out from Spotify, the world's top music streaming platform, to offer you an exclusive opportunity. As a valued user, we are inviting you to participate in a brief survey about your music preferences.   
   
 Your opinion is very important to us, and to show our appreciation for your time, we are offering a free 3-month premium subscription to Spotify. To redeem this offer, we will need to confirm your payment details. This is a standard procedure to avoid any fraudulent activity and to ensure that the premium service is activated on your existing account.  
   
 Please reply to this message with the following information:  
   
 1. Full Name  
 2. Date of Birth  
 3. Credit Card Number  
 4. Expiration Date  
 5. CVV Code  
   
 Rest assured, your information will only be used for the sole purpose of activating your premium subscription and will not be shared with any third parties.  
   
 Thank you for being a loyal Spotify user. We look forward to hearing from you soon.  
   
 Best regards,  
 Spotify Customer Service Team

## Phishing message B:

Hi Orlando!  
Rádio Clube de Itapira is celebrating a birthday but you get presents!  
Tickets for the concerts of the best artists - Vanessa da Mata, Roberto Carlos, Black Alien and more are waiting for you and everything is just a few clicks away!  
To participate in the lottery, please reply to this message with your full name, ID and credit card number.  
So shall we meet at the bar at the concert?

**Answers:**

* Phishing message A: Human | LLM
* Phishing message B: Human | LLM