

Activist

You will be shown a series of message and you should classify them as SPAM or not. Note that personal fields will have this generic format (your field) or for senders (field) and official logo images are to be considered as actual/official values. In order to avoid any copyright violation names and images were altered.

Let's assume these are actual companies:

- **Pear:** Is a company that sells yMobile smartphones. Official website : pear.co.uk
- **PayFriend:** Is a company that helps with online payments. Official website: payfriend.com. Email address: payfriends@payfriend.com
- **Recruitron:** Is a recruiter company. Official website: recruitron-business.com. Email address: contact@recruitron-business.com
- **PhotoZ:** Is an art sharing website. Official website: photo-z.art . Email address: photoZ-contact@photo-z.art
- **ReZap:** Networking website like LinkedIn. Official website: rezap-future.com. Email address: contact@rezap-future.com
- **StarBank:** Is a bank's website. Official website: starbanking.co.uk Email address: info@starbanking-contact.co.uk

1.

From: marketing@pear-ymobil3.co.uk

Subject: ONCE IN A LIFETIME OFFER !

Get the new 32GB yMobile for £ 200 ONLY !

BUY NOW

CLICK [HERE](#) !

2.

From: payfriend@pay-friend.co.uk

Subject: View your recent transactions now !



Email address: (your email address)

Date

15th March 2015

Hello valued customer,

It's nearly St Patrick's Day and it's time to get into the festive mood and you can securely do this with Pay Friend's support.

Remember, to view your activity, just log in to your account.

[Log in](#)

3.

From: Alaska@ contact@recruitron-business.com

Subject: Recruitron Your Profile

Hi (Your name)

I represent a company here in Glasgow that is looking for a (your career) graduate position and your profile stood out to me.

If you are interested, please let me know.

Kind Regards

4.

From: PhotoZ <sponsor@e-offer.biz>

Subject:

Hi,

You have message from Linda on PhotoZ:

I saw your photo & thought it'd be great for the Nature challenge.
Enter it now OR take a new photo by Sunday.

- Linda

Go to your [inbox](#)

5.

From: ReZap<rezap@r3-zap.co.uk>

Subject: People are viewing your profile

Attachments: (offer.exe)

Hi (username),

Your profile has been viewed by these users:

- Marie K. Bonnet

- Bianca Delano

- Magnolia Fisher

- Gia G.

Gia G. has endorsed you : + Java + C

G E T TH3 NEW 1PH0N3 H

ERE 4 17o P ou n ds

Remove the captial R and spaces:

ofRR ferRR-RRiRRp ho neRR6RR.cR Rom

6.

From: StarBanking<info@starbanking-contact.co.uk>

Subject: Account plans are being changed

Is this e-mail displayed correctly? View it in your browser. Your partial postcode is : (your post code)



Dear (Your Full Name),

We are glad to inform you that your account has been upgrade to an (...) account due to improvements in our services.

[Log in](#) to your account to view more details about your new plan.

Kind Regards,

StarBank (...) Team

Correct answers:

1. Yes – Talking about once in a lifetime opportunity (urging user), using all capitals, fake or suspicious sender address and suspicious link.
2. Yes – This is a replica of an e-mail sent by PayFriend if you look at the sender address it looks like their address but the real one is payfriends@payfriend.com . The log in link does not redirect to their official website. Also refers to you as dear valued customer
3. No – It's not a spam e-mail. The e-mail does not urge the reader to take immediate action. The company's site seems to be legit. It uses your real name and looks like a job offer of your expertise.
4. Yes- Sender impersonating PhotoZ website. Link redirects to their website's log in page. Empty subject
5. Yes – Suspicious attachment. Suspicious spaces, line breaks and characters. Hidden text
6. No - Uses your full name, links send you to their official website and do not request any information from you

Reflector

The purpose of this exercise is to classify Spam/Phishing e-mails and talk about the reasons behind the classification.

Let's assume these are actual companies:

- **Pear:** Is a company that sells yMobile smartphones. Official website : pear.co.uk
- **PayFriend:** Is a company that helps with online payments. Official website: payfriend.com. Email address: payfriends@payfriend.com

1. E-mail example

From:marketing@pear-ymobil3.co.uk 1. **Suspicious address/ Does not look like pear's e-mail address**

Subject: ONCE IN A LIFETIME OFFER ! 2. **Excessive use of caps - Limited offer/Immediate action**

Get the new 32GB yMobile for £ 200 ONLY !

BUY NOW

CLICK HERE ! 3. **Suspicious website link / Not apple's website**

Spam

This message has been classified as spam/phishing e-mail because:

- Suspicious sender impersonating Apple
- All caps in subject and use of caps throughout the e-mail should be considered as suspicious when the use is excessive
- The link that the e-mail uses is not an official apple website and it is most likely a phishing website where details are recorded and/or any transaction done is not associated with the purchase of any product.

This might a simple example but spam/phishing e-mails/websites can be difficult to identify at a first glance:

- Legitimate website and an illegitimate pop up window
 - Illegitimate website can be identical to the legitimate one
-

2. E-mail example 2

From: payfriend@pay-friend.co.uk **E-mail address looks legitimate however the real paypal address is <payfriends@payfriend.com>**

Subject: View your recent transactions now.



Email address: (your email address)	Date
	15th Mar 2015

Hello valued customer, **Generic greeting**

Hello valued customer,

It's nearly St Patrick's Day and it's time to get into the festive mood and you can securely do this with Pay Friend's support.

Remember, to view your activity, just log in to your account.

Log in **Suspicious link. The address does not look like PayPal's website address**


Spam

This message has been classified as spam/phishing e-mail because:

- Suspicious sender impersonating PayFriend
- Generic greeting. Use of "valued customer"
- The link that the e-mail uses is not an official PayFriend website and it is most likely a phishing website where details are recorded and/or any transaction done is not associated with the purchase of any product.

3. Website example

← → ↻ 🏠 cybersecurityproject.info.com/accounts/login/ **URL does not seem right**

 **Low quality**

TRAINING PORTAL [🏠 Home](#)

Log In

Log in with your **Training Portal** account:

Username

Password

☐ **Remember Me**

[Forgot Password?](#) **Log In**

If you do not have an account, please [sign up](#)

Phishing website

This website has been classified as a phishing website because:

- Website address close to legitimate website but it's different
- Images are of lower quality
- Theme / Website's appearance is a bit off
- Log in might accept any information -> no validation just recording details

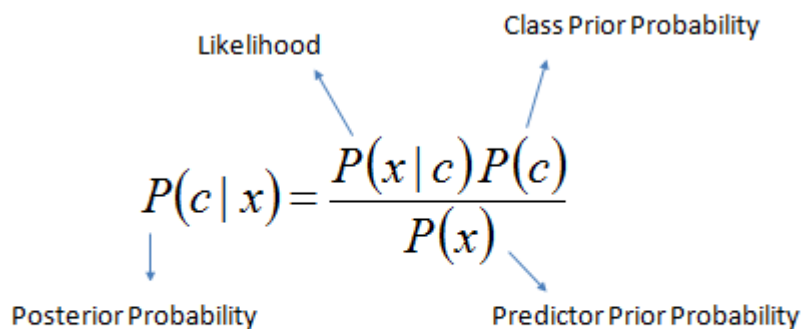
Possibly used to gather account information and gain access to accounts of that website.

Note: Phishing websites can look identical to a legitimate site and the signs are not that obvious. Some hints for identifying phishing sites are to look for HTTPS (secure access) in the address (for banking services). However beware phishing websites can also use HTTPS.

Theorist

Spam filters use text categorization to sort messages. One of the algorithms used is the Naive Bayes algorithm in order to categorize messages into spam and non-spam categories. The algorithm calculates probabilities based on a series of messages that have been manually annotated. When a message is processed, all the words are taken and analysed against the probabilities of them appearing in spam message versus them appearing in non-spam messages. If a message has an overall spam probability higher than the non-spam probability, then the message is classified as spam and then dealt with accordingly by the software.

Naive Bays formula [Source: http://www.saedsayad.com/naive_bayesian.htm]

$$P(c | x) = \frac{P(x | c)P(c)}{P(x)}$$


The diagram illustrates the components of the Naive Bayes formula. It shows the equation $P(c | x) = \frac{P(x | c)P(c)}{P(x)}$ with four labels and arrows pointing to specific parts: 'Likelihood' points to $P(x | c)$, 'Class Prior Probability' points to $P(c)$, 'Posterior Probability' points to $P(c | x)$, and 'Predictor Prior Probability' points to $P(x)$.

$$P(c | X) = P(x_1 | c) \times P(x_2 | c) \times \dots \times P(x_n | c) \times P(c)$$

- $P(c|x)$ is the posterior probability of *class (target)* given *predictor (attribute)*.
- $P(c)$ is the prior probability of *class*.
- $P(x|c)$ is the likelihood which is the probability of *predictor* given *class*.
- $P(x)$ is the prior probability of *predictor*.

Example

Given 20 normal e-mails and 20 spam e-mails, here is a table of how many times 'Free', 'Cat' occurred

Word	-		
-	Type	Spam	Not Spam
Free	-	15	5
Cat	-	1	10

Taking the word 'Free'.

$$P(\text{Spam})=0.5$$

$$P(\text{Not Spam})=0.5$$

$$P(\text{Free}|\text{Not Spam}) = 5/40 = 0.125$$

$$P(\text{Free}|\text{Spam}) = 15/40 = 0.375$$

$$P(\text{Not Spam}|\text{Free}) = P(\text{Not Spam}) * P(\text{Free}|\text{Not Spam}) = 0.5 * 0.125 = 0.06$$

$$P(\text{Spam}|\text{Free}) = P(\text{Spam}) * P(\text{Free}|\text{Spam}) = 0.5 * 0.375 = 0.19$$

This word Free will be classified as a spam related word because there is a higher probability that it belongs to spam.

Calculate probabilities for the the word Cat.

Bibliography

1.

Text categorization.Yiming Yang and Thorsten Joachims (2008), Scholarpedia. Available at: <http://www.scholarpedia.org/arti>

2. Choochart Haruechaiyasak.(2008). A Tutorial on Naive Bayes Classification . Available at: <http://suanpalm3.kmutnb.ac.th/teacher/FileDL/choochart82255418560.pdf>
3. Naive Bayesian, saedsayad. Available at: http://www.saedsayad.com/naive_bayesian.htm
4. Karen Rubin.(2012). The Ultimate List of Email Spam Trigger Words. Available at: <http://blog.hubspot.com/blog/tabid/6307/bid/30684/The-Ultimate-List-of-Email-SPAM-Trigger-Words.aspx>
5. Dell Zhang. (2006). An Example of Text Classification with Naïve Bayes. Available at: http://www.dcs.bbk.ac.uk/~dell/teaching/ir/examples/nb_example.pdf
6. Dell Software, Naive Bayes Classifier. Available at: <http://www.statsoft.com/textbook/naive-bayes-classifier>

Pragmatist

There various online tools that check the deliverability of a message by assigning a quality score. This score determines how likely it is to be marked as spam based on the subject, content and the sender of the message.

ISnotSpam and Mail-Tester helps determine if a message will be classified as spam by forwarding the email to a given address and viewing the report

ISnotSPAM

<http://isnotspam.com/>

mailtester

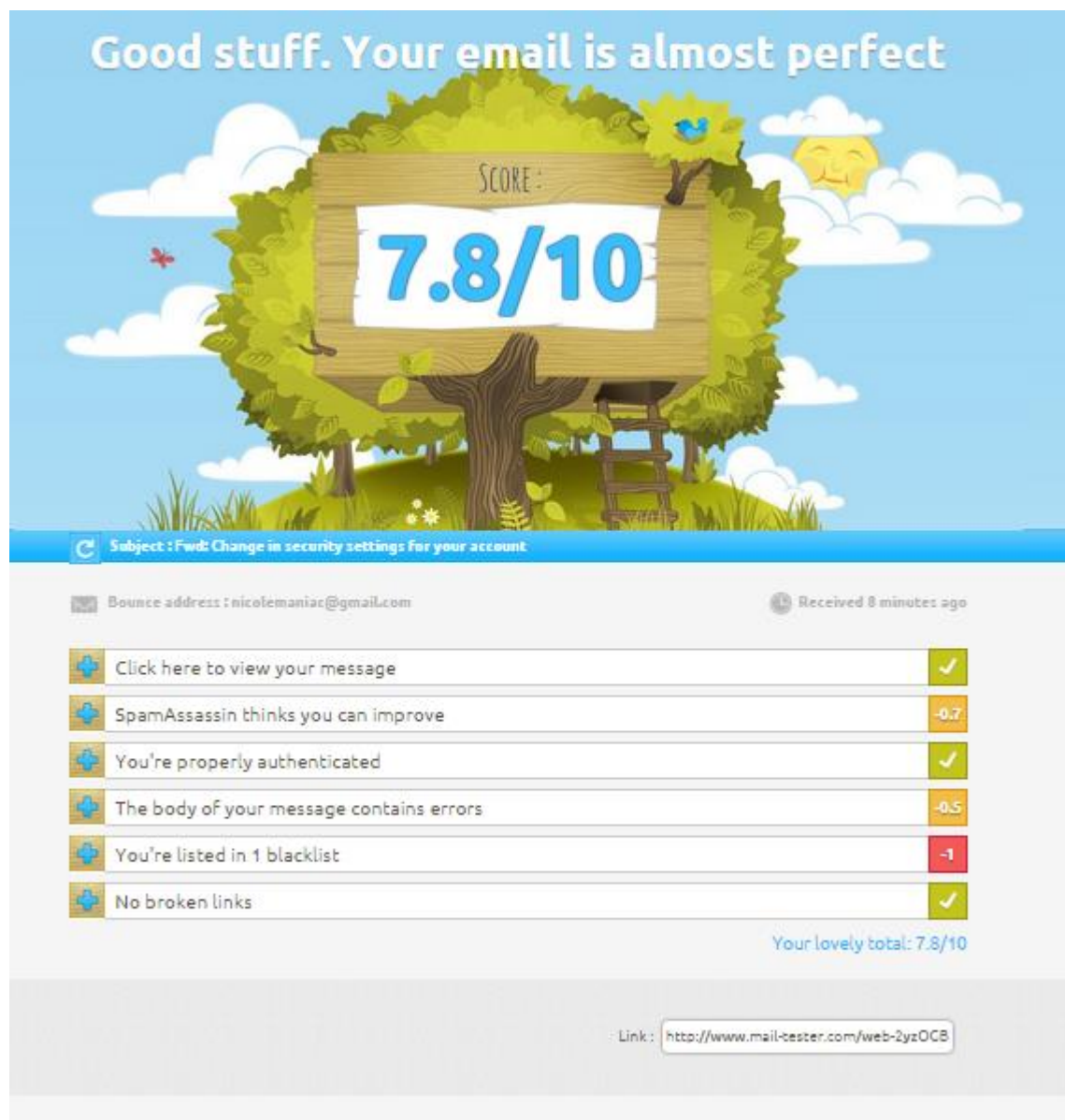
<https://www.mail-tester.com/>

Example

An e-mail sent to both ISnotSPAM and mail-tester in order to check the quality score.

These were the results:

1. Mail-Tester



Example from the mail-tester websites (<https://www.mail-tester.com/>)

2. IShotSPAM

ISnotSPAM Latest Report Details

[Delete this report](#)

This message is an automatic response from isNOTspam's authentication verifier service. The service performs a check of various sender authentication mechanisms. It is provided free of charge, in the hope that it will help. If not officially supported, we welcome any feedback you may have at .

Thank you for using isNOTspam.

The isNOTspam team

=====
Summary of Results
=====

SPF Check : pass
Sender-ID Check : pass
DomainKeys Check : neutral
DKIM Check : pass
SpamAssassin Check : ham (non-spam)

=====
Details:
=====

HELO hostname: mail-ig0-f180.google.com
Source IP: 209.85.213.180

Example from the mail-tester websites (<http://isnotspam.com/>)

Spam detection

In order to activate spam detection you can use a generic test string called GTUBE that is used to test/trigger spam filters.

XJS*C4JDBQADN1.NSBN3*2IDNEN*GTUBE-STANDARD-ANTI-UBE-TEST-EMAIL*C.34X

References:

<http://en.wikipedia.org/wiki/GTUBE>