**Phishing Emails**

For this project we decided to have 5 categories to determine if an email is a phishing email or not. The categories we have decided on are: urgency words, confirmation words, The Email Address itself, the greeting, and URLs that were attached. Each category is not equally weighted as we believed that some words of an email are more telling if an email is a phishing email then others.

Urgency words have a weight of 5/20. If an email is continually using words that pressure the reader to do something immediately or use seductive wording for the reader to change their lives. This exception is thrown. This exception has higher score then most the others because we reasoned that most normal emails are not using words to push the reader for an immediate response.

Confirmation words have a weight of 6/20. This exception is all about an email asking for information from the user. Whether it be bank information, or the email asking for the user’s social security number this exception will be thrown and is our heaviest scale. The reason being is that almost any time someone is asking for this personal information over an email it is a phishing attack. However, we did not make it heavy enough for an email to fail off just this exception, on the rare case that it is not malicious.

The Email address has a weight of 5/20. There are three main parts of the Email address that our checker is looking for. The first is to check the ID to determine if it has any uncommon elements to it. For instance, if an email has a large amount of numbers in the string it will be flagged with 2 of the 5 weight points. Next is to check the Domain. The first check is to make sure there are no numbers in the domain. If there are numbers, the program will be flagged and another 2 of the 5 weight points will be added. Lastly, the domain is checked to see if it is a common domain or not. For example, if a domain is @sketchy.com this will be flagged and will add another 1 of the 5 weight points. We believed the email domain to have a higher weight as well due to being one of the easier ways that we were able to detect phishing attacks.

Greetings only have a weight of 2/20. Greetings are the how the email addresses the user. If the greeting is unprofessional or impersonal, the program is flagged and will add the 2 points to the score. We thought this one was valid due to many phishing emails being rather generic in their greetings, or unprofessional at times. It does not have a heavy weight however due to the fact there are many who do not write good greetings and it is not always malicious.

URLs also only have a weight of 2/20. URLs that are sent to the user automatically flags the program and adds the two points. Therefore, the weight is so low for URLs. There are many who send URL tags to individuals through Email and are not seeking to harm the user. However, a URL is still suspicious when it is combined with the other elements listed above so that is why it is flagged with true or false.

No one category will automatically mark the email as a Phishing attack as we are giving the benefit of the doubt that some emails will just be poorly written or have odd things in it but not be harmful.

To improve upon this Email checker, there are 4 additional tools that could be added to the future to make this an even more effective tool to be used. The first is a spell checker. This would be a low weighted item but could still prove useful. Next is a grammar checker. Many Phishing emails have terrible grammar and although not a sure-fire way to find a phishing attack can help with the process. Next is location detection of where the email was sent. This one would be the most difficult to implement and would take a lot of resources, however if an email was able to be traced to a place that is not a normal place for the user to be receiving emails from it could effectively fish out a phishing attack. Lastly is to install an API for one of the URL checkers that were mentioned in class. One of these could check if any of the URL tags that were sent were illegitimate and could be harmful to the user.