Team name: BEAMING
Team members: Dilay Ozkan, Nghia Nguyen, Benjamin Torres, Malavika Mampally, Egemen Elver
Date: 03/29/2024

Team roles for this report (write down name):

Facilitator(s): Dilay Ozkan, Egemen Elver

Recorder(s): Nghia Nguyen

Deliverer(s): Nghia Nguyen

Planner(s): Benjamin Torres, Malavika Mampally

Team Contact: Dilay Ozkan (dilay@unc.edu)

- **Describe briefly what the main goal of your team is (so the peer reviewer has some context). E.g. we are working on image classification for blah de blah. Our goal is blah de blah etc. In the initial part of the semester before your proposal it is ok to put down "we are still coming up with ideas on team project".**

Our team's goal is to apply the methods we learn from classes and external sources to handle the problem of identifying phishing web pages using the dataset publicly available dataset of Hannouse and Yahiouche (2021) by developing and improving some classification models.

I. **What was done during the report period regarding the project:**

So far, we finalized the dataset and the problem. We decided to work on a classification problem regarding whether a website is phishing or legitimate based on several features extracted from the URLs as well as the website's content. We also started working on some ordinary methods, including statistical methods and tree-based methods.

Specifically, Dilay and Nghia have been working on logistic regression and decision tree methods (CART). Benjamin has been working on random forest and gradient-boosting tree methods. Egemen has been working on Support Vector Machine methods. Malavika has been working on the K-nearest neighbors' method. It's surprising that basic methods provide good accuracies (above 90%).

II. What were obstacles faced if any in working on the project?

- The first obstacle is time limitation. We have to take other classes, do research, and teach/ do TA, so we can't spend much time on the project.
- Some of us are not that familiar with Python and Google Colaboratory compared to R, so it may take us more time to get familiar with the commands and structures in Python.
- Some models take time to implement. For example, SVM with radial kernel takes much time so we will leave it to the next biweekly report period.

III. What is the plan for the next reporting period including what each team member is planning to work on. Describe goals and potential timelines.

1. We will implement some other traditional methods, including Naive-Bayes classification, LDA, and QDA, as well as advanced methods, including convolutional neural networks (CNN), long-short term memory (LSTM), and multilayer perceptron (MLP).
2. We will compare the performance of the considered methods and try to give some comments about their performance. Based on these results, we can validate the efficiency of the methods we learn from classes and from external resources.
3. We will also start working on the report and preparing the presentation