

AI SECURITY BOOTCAMP

A PRACTICAL GUIDE ON SECURING AI SYSTEMS

📄 We are looking forward to welcoming you to our upcoming training on Adversarial Machine Learning. This interactive session will dive into using Foolbox to test and analyze the robustness of machine learning models against adversarial attacks. Below are important details and pre-requisites to ensure you are fully prepared for the workshop.

This pre-training document is designed to help you familiarise with the environment and some tools that we will be using.

There is not need for you to do any preparation other than to make sure you are familiar with these platforms and tools. Particularly, how to access them.

We'll be using the following for our training:

- Google Colab
- GitHub
- Python (all code will be provided to get started)
- Tensorflow
- Attack frameworks (eg. Foolbox, TextAttack etc)
- Stable Internet and Chrome Browser (highly recommended, Colab works best with Chrome)

Workshop Platform: Google Colab

- Access Requirements: Google Colab is a free cloud service hosted by Google to encourage machine learning education and development. To access Google Colab, you will need a basic Gmail account. If you do not already have one, please create a Gmail account prior to the workshop at <https://accounts.google.com/SignUp>.
- Features of Google Colab: Colab allows you to write and execute Python in your browser with:
 - No configuration required
 - Access to free GPUs and TPUs
 - Easy sharing capabilities

Workshop Materials and Code:

- GitHub Repository: All workshop materials, including code snippets and pre-trained models, will be available in the trainer's GitHub repository. We will provide access instructions at the beginning of the session, allowing you to clone or download all necessary resources directly into your Google Colab environment.
- Repository URL: The specific URL for the repository will be shared at the workshop.

Pre-Workshop Preparation:

- Familiarize Yourself with Python: A basic understanding of Python programming is essential for this workshop. If you are new to Python, we recommend reviewing basic syntax and concepts.
- Review Basic Machine Learning Concepts: While deep knowledge of machine learning is not required, familiarity with general concepts will be beneficial.

You can get started with these articles:

<https://malienist.medium.com/ai-cyber-security-bootcamp-setting-up-the-environment-8919300b15cc>

<https://malienist.medium.com/ai-security-the-concept-of-underfitting-1d491c65e108>

<https://malienist.medium.com/ai-security-the-concept-of-overfitting-addef0071aa4>

- Set Up Your Workspace: Ensure you have a stable internet connection and access to a web browser (Google Chrome or Firefox recommended) for the best experience with Google Colab.

Workshop Agenda:

- Introduction to Adversarial Machine Learning
- Overview of Google Colab
- Training AI Models and working with common datasets
- Hands-on Attacks
- Visualizing and Analyzing Results
- Q&A and Discussion


We encourage you to arrive on time and ready for action! Should you have any questions ahead of the workshop, please do not hesitate to reach out.

Looking forward to a productive session!

Best regards,

Vishal Thakur

<https://www.linkedin.com/in/malienist/>

 Workshop Date: 9 June 2024

Time: 13:30 – 17:30