What is this about?

It is about How can we automate the recommendation of attack steps/sequences? How to quantify chances of success and impact of each step given preferences (e.g., stealth vs impact)?

In order to automate. Needs to have information first. Use EdRange as example and vulnerability scans such as OpenVAS for automation.

OpenVAS output the results as pdf and/or xml. It seems like xml is better to convert into machine understand language.

Use python for machine learning and xml converting.

For some unknown reason outlook unable to send xml files. Pdf files is fine. So, I will use GitHub to get the scans results to my real computer instead of the VM. Which does not work so I just use mega instead.

Note: the free version of OpenVAS is only useable in VM.

Got a new PC. Setting everything up again….

I9-13th gen, rtx 4070!!!! DDR5 RAMMMMM SPPEEEEDDDDD POWERRRRRRRRRR!!!#E@\*&#\*(!)&#)!\*@&#.

At least it runs very smooth.

Found out that MITRE ATT&CK actually have a GitHub repo that can request data from the website. [mitre-attack/mitreattack-python: A python module for working with ATT&CK](https://github.com/mitre-attack/mitreattack-python/)

Might be useful.

Script is disable by default in windows. Nice. [execution of scripts is disabled how do you enable? - Microsoft Community](https://answers.microsoft.com/en-us/windows/forum/all/execution-of-scripts-is-disabled-how-do-you-enable/e19d41b2-ab61-e011-8dfc-68b599b31bf5)

# 1

Watched a video about machine learning with python: [(1) Python Machine Learning Tutorial (Data Science) - YouTube](https://www.youtube.com/watch?v=7eh4d6sabA0&t=1861s)

Steps from the video:

Import Data

Clean the Data

Split the Data into Training /test sets

Create a Model

Train the Model

Make Predictions

Evaluate and improve

Import the Data! I need more data!!!!!!!

Found a stack overflow post about how to extract the text from xml with python.

[Extracting text from XML using python - Stack Overflow](https://stackoverflow.com/questions/7691514/extracting-text-from-xml-using-python)

But the code seems old and something is wrong.

return parser.close()

~~~~~~~~~~~~^^

xml.etree.ElementTree.ParseError: syntax error: line 1, column 0

Decided to use official documentation:

[xml.etree.ElementTree — The ElementTree XML API — Python 3.13.0 documentation](https://docs.python.org/3/library/xml.etree.elementtree.html)

Successfully extract the data from the scans results.

[python - How to print formatted string in Python3? - Stack Overflow](https://stackoverflow.com/questions/26862773/how-to-print-formatted-string-in-python3)

I changed my mind. I got bored of this machine learning stuff. I realize I am actually interested into LLM instead of this.

Plus, I got my new PC with RTX 4070 12GB. Should be enough to fine tune a small model. Only one way to find out.

Introducing Hugging Face. Basically, a GitHub for LLM.

Like anyone would do. I put my eyes at the strongest open-source model out there. nvidia/Llama-3.1-Nemotron-70B-Instruct-HF seems to be nice. It appears to be match to GPT-4o in some degree. As I check its requirements. I realized I was too native. **It required at least 2 or more 80GB GPUs**. And my is 12GB…. Not even close…

So ultimately, I come with 2 options. I can either get a pre-trained models and fine-tune it to match this project. Or I can make a model from scratch.

I choose to get a pre-trained model. Simply because my lack of knowledge of how all these work in foundation. Thus, I would learn fine tuning first alone with all the LLM installations and training.

I choose Mistral (7B). cause why not. Also, because LLaMA required me to put down my contact information…

Found a webui for the LLM. Used one for stable diffusion before so this is nice.

[oobabooga/text-generation-webui: A Gradio web UI for Large Language Models.](https://github.com/oobabooga/text-generation-webui)

Turns out that webui has built in downloader from hugging face. That’s nice. I can just provide the link and it will auto download and deploy.

I try [nvidia/Llama-3.1-Nemotron-70B-Instruct-HF · Hugging Face](https://huggingface.co/nvidia/Llama-3.1-Nemotron-70B-Instruct-HF),

[mistralai/Mistral-7B-Instruct-v0.2 · Hugging Face](https://huggingface.co/mistralai/Mistral-7B-Instruct-v0.2).

While they download, which takes forever. I watched a few more video about how to fine tune LLM.

[EASIEST Way to Fine-Tune a LLM and Use It With Ollama](https://www.youtube.com/watch?v=pxhkDaKzBaY&t=125s)

Then I realized the webui also has a built-in trainer with tutorial. Which is nice.

[05 ‐ Training Tab · oobabooga/text-generation-webui Wiki](https://github.com/oobabooga/text-generation-webui/wiki/05-%E2%80%90-Training-Tab)

Funny how it actually supports Raw text file training. Which means I could probably use this to get results pretty fast.

Now I just need to gather my datasets. Which is scans results from OpenVAS, all the MITRE ATT&CK thingy and Metasploit’s documentations. See if this work.

WHHHHHHHHHHHHHHHHHHHHHHHHHHHHYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY IIIIIIIIIIIIIIIIIIIIIIISSSSSSSSSSSSSSSSSSSS ITTTTTTTTTTTTTTTT TAKKING G SOOOOO LONNGGGGGG. It has beeeeeeeeeeeeeen 2 hours since I clicked the install buttommmodmomdoawmdoanocunsadoc

I literally have internet speed of 20MB perrrrr secondddddddAHHHAOUHOUWDHA

It’s not even at the install process yet. It is still downloading. GOD.

At least the script that turns Report.xml file into short and simple text for LLM to read worked fine. “extract\_vulnerabilities.py”

So, without surprise. This 2-3 hour downloaded model straight up crash when I try to load it. Well, back to Mistral. Eyy, turns out I need to be authorized in order to download models. Need to figure out how to login. Well, I just download everything manually and put them into the model folder.

In the meantime, of all the install. This video is the by far the best.

[host ALL your AI locally](https://www.youtube.com/watch?v=Wjrdr0NU4Sk)

I probably will start everything over just to follow this guide.

Sooooo,

Mistral 7B worked. But very slow.

Welp I am just going to restart everything and follow this video guide.

[host ALL your AI locally](https://www.youtube.com/watch?v=Wjrdr0NU4Sk)

Install Ollama! Need to enable WSL for win 11 first.

Successfully enable wsl and got ubuntu!

Found out that Nvidia cuda driver was not installed. Maybe the reason why the webui was running so slow.

Going to test it after cuda is install. Never mind. Just realized I deleted everything. I gotta be patience. Damn

Installing the llama 3.

IT RUNN SO SMOOTH. SO, THIS IS what 4070 SUPPOSE TO BEEEE.

Time to add the model into a webui so it looks better.

It also has text guides nice! [The NetworkChuck Academy](https://academy.networkchuck.com/products/youtube-videos/categories/2155282450/posts/2177513911)

A new webui call open-webui!

All setup and complete. Run smooth.

Now I just need to figure out how to fine-tune it for my needs.

[EASIEST Way to Fine-Tune LLAMA-3.2 and Run it in Ollama](https://www.youtube.com/watch?v=YZW3pkIR-YE)