WEAPON DETECTION USING NEURAL NETWORK CLASSIFICATION

DATA SCIENCE PRESENTATION BY: MARSHALL CYRUS

PROBLEM STATEMENT

Image object detection is a prominent and emergent area of study.

Photographs and video are available in abundance all around the world.

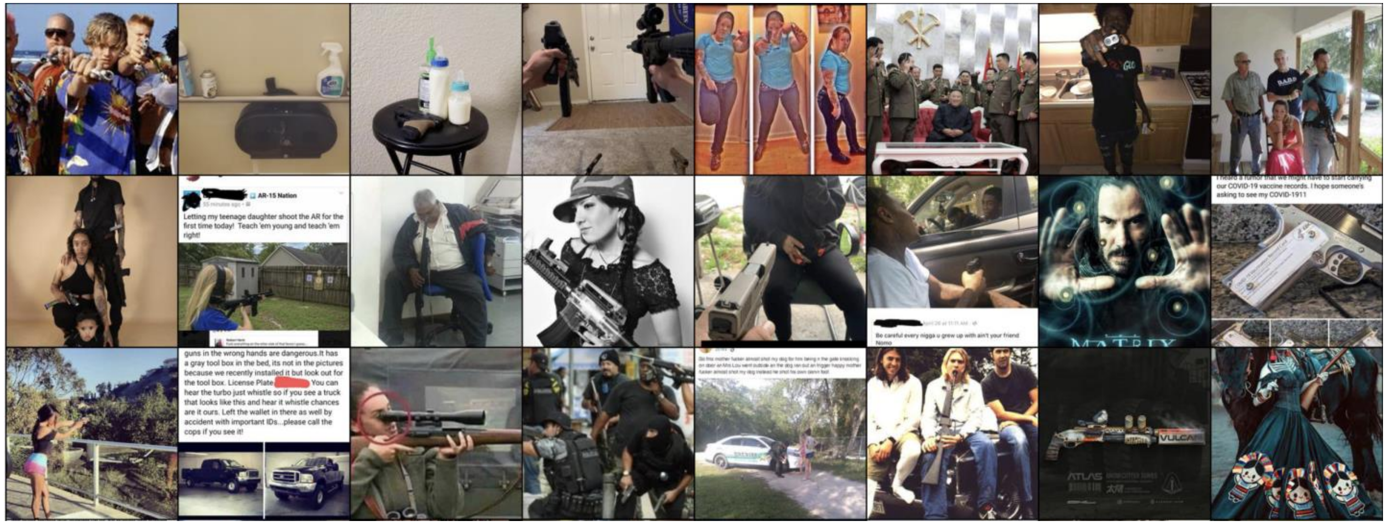
As a third party contractor for the government, I am tasked with being able to identify photos that contain a person holding a firearm so that public surveillance cameras can give an alert to law enforcement.

BACKGROUND

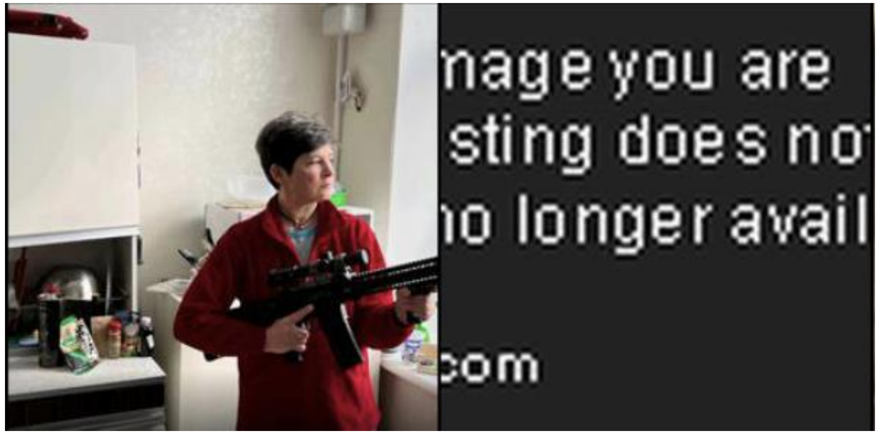
- An automatic weapon detection system can provide the early detection of potentially violent situations that is of paramount importance for citizens security.
- One way to prevent these situations is by detecting the presence of dangerous objects in surveillance videos.
- Deep Learning techniques based on Convolutional Neural Networks can be trained to detect this type of object.

DATASET SELECTION

- OD weapon master set (https://github.com/ari-dasci/OD-WeaponDetection)
- Reddit
  - Idiotswithguns, ChicksWithGuns, TheWayWeWere, Portraits, Guns, Military, Pics, Shooting, Airsoft, Shotguns
- Custom set with Iphone and Google



DATA COLLECTION DIFFICULTIES



- Over 45 thousand images in varying setting and sources.
- Finding applicable photos of an individual holding a firearm (Scrape, download, resize/crop).
- Manual examination using an image grid.

WORKING PROCESS – COLAB AND GPU INTEGRATION

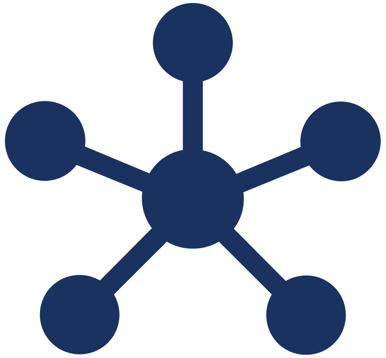
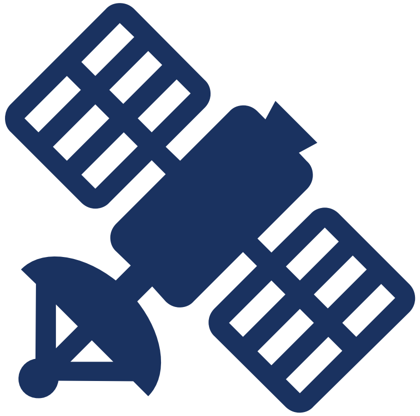
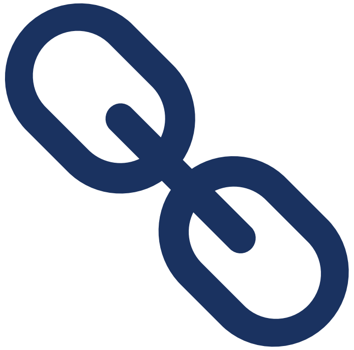


Image Collection



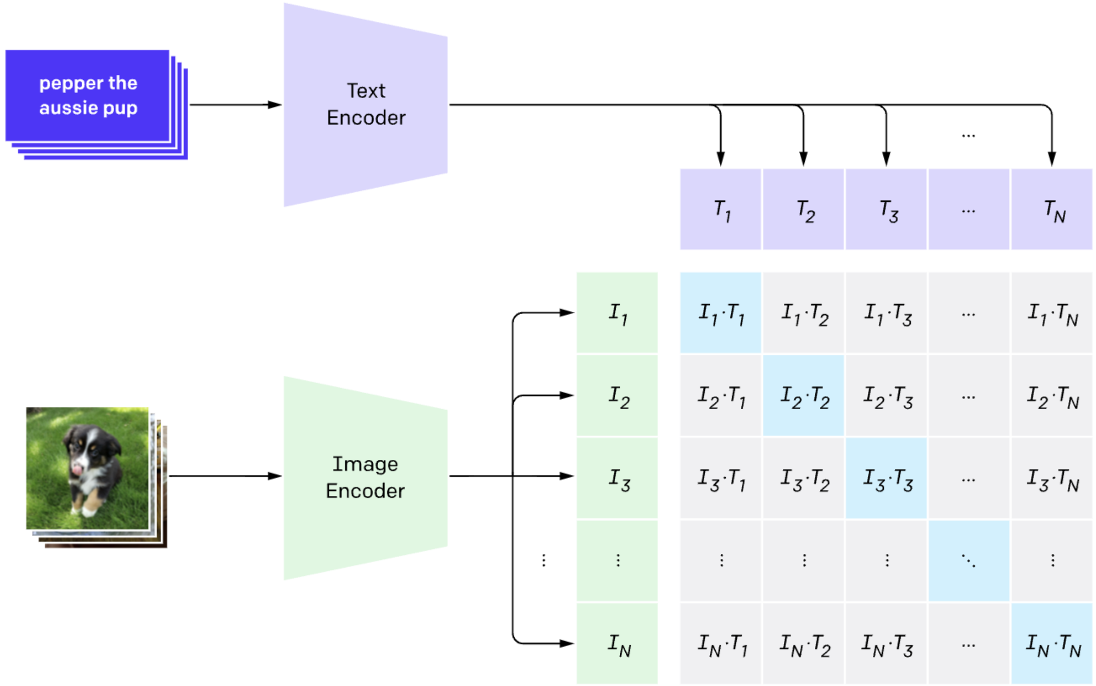
Preprocess & CLIP



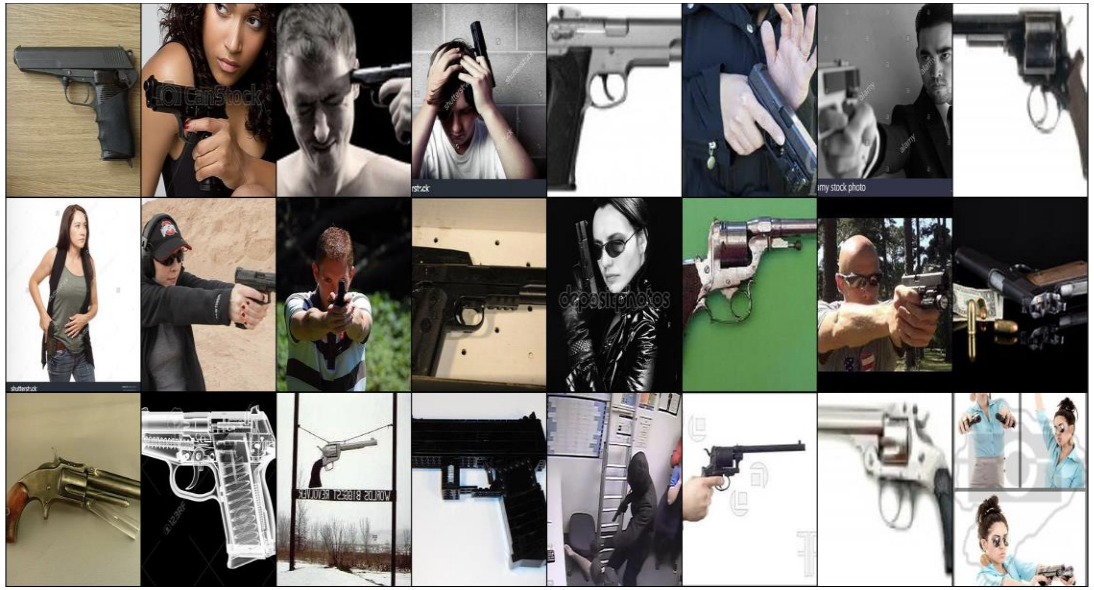
CNN Evaluation

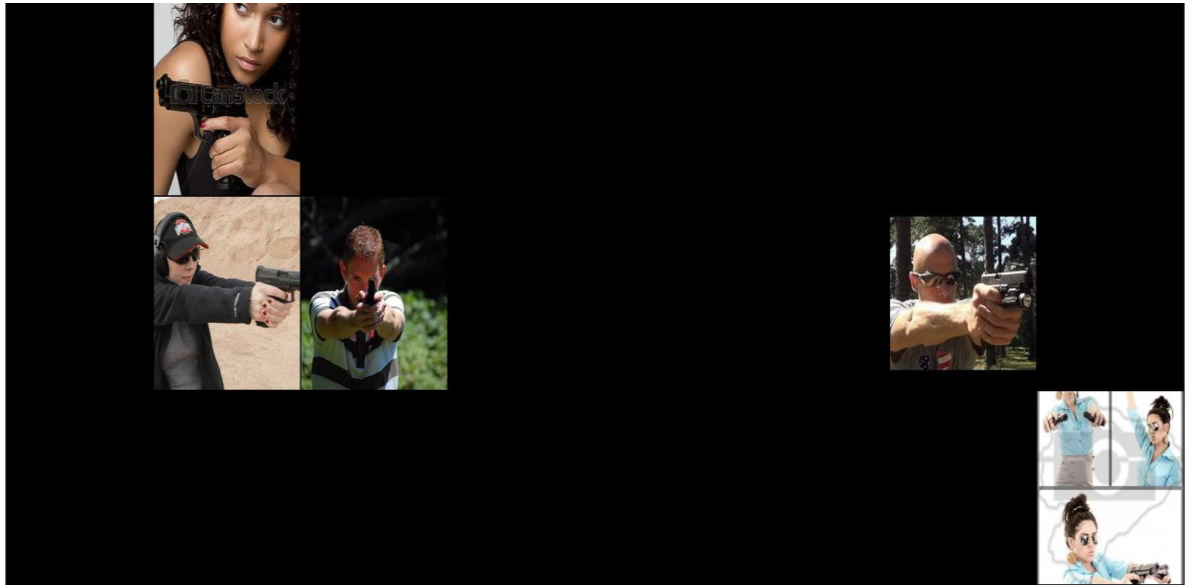
CONTRASTIVE LANGUAGE IMAGE PROCESSING

- Learns visual concepts from natural language supervision.
- Highly efficient and flexible, trained on 400million images with 256 GPU's over 2 weeks.
- However, cannot learn fine grain details



CLIP EXAMPLE

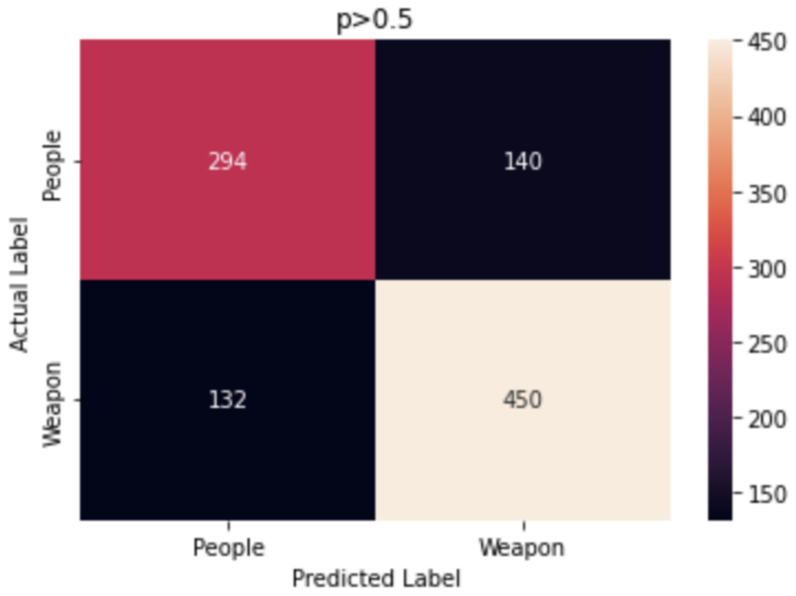


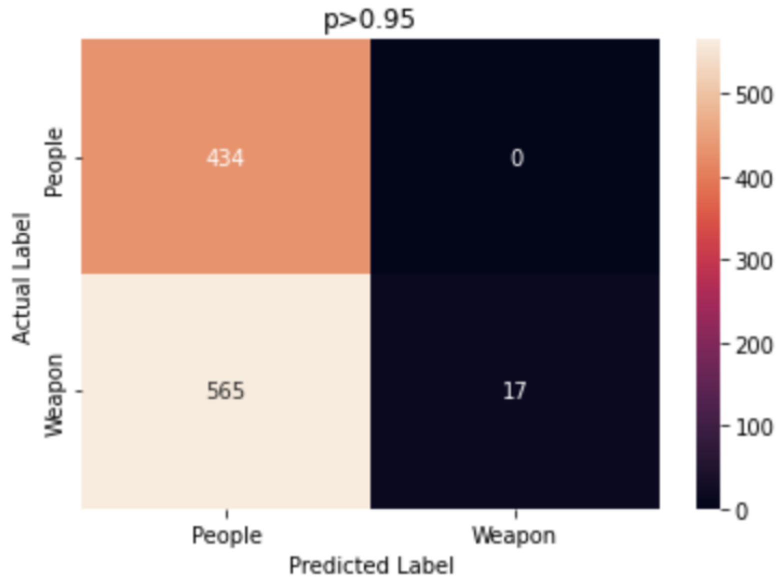


```
caption1 = 'A photo containing a person'
caption2= 'A photo not containing a person'
g1,g2_e= examine_grid(images,caption1,caption2,nrows=8,prob_cutoff=0.78)
```

INITIAL CNN MODELING

- 67,369,281 parameters - val_accuracy: 0.7037
  - Increase Model Size and Add Batchnorm and Dropout
  - Add Data Augmentation. Increase Network Capacity
- Confidence interval cutoff

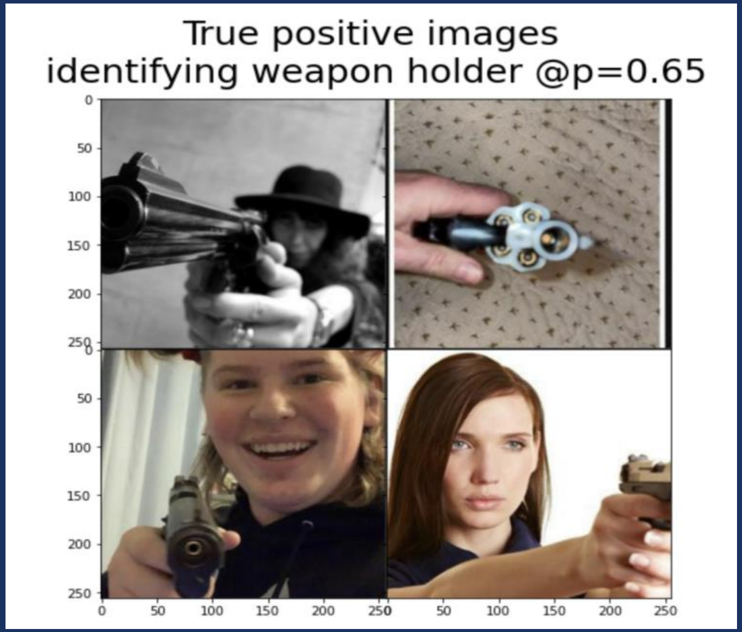


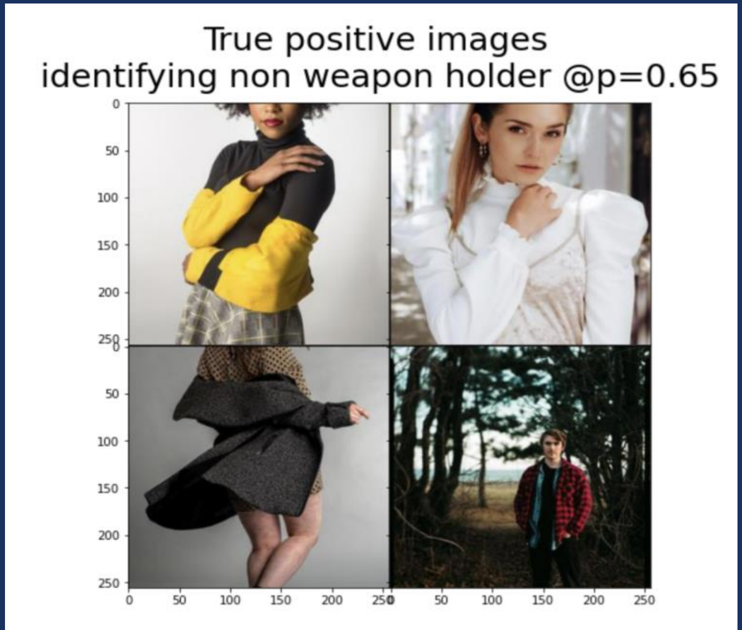


MODELS USING TRANSFER LEARNING

METRICS - TRUE POSITIVES, FALSE POSITIVES, TRUE NEGATIVES, FALSE NEGATIVES, BINARY ACCURACY, PRECISION, RECALL, AUC, PRC

val_accuracy: 0.9357 - val_precision: 0.9608 - val_recall: 0.9304 - val_auc: 0.9805 - val_prc: 0.9877





ADVERSE DATA

- Pushing model limits using images that show a human holding an item, however not a weapon.
- Guitars, boxes, golf clubs etc..

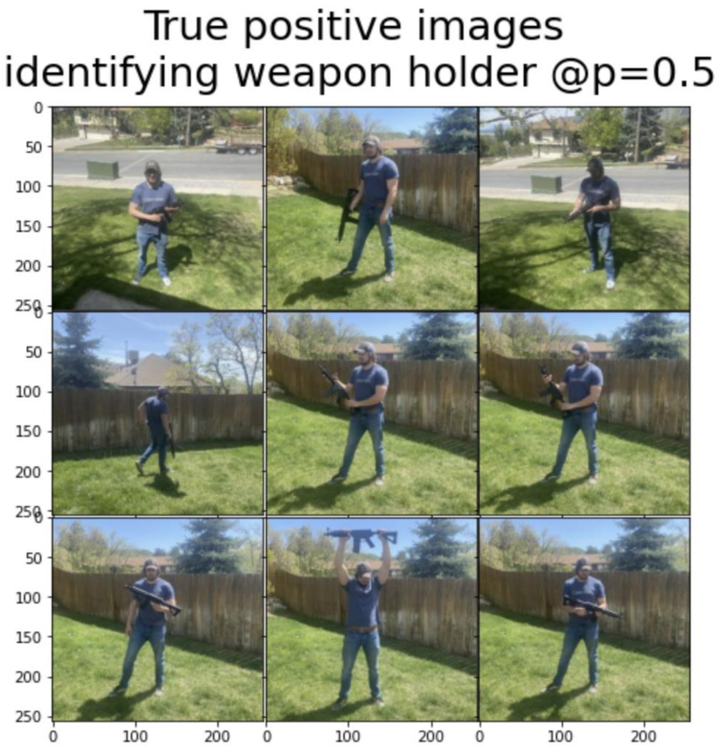
	loss	tp	fp	tn	fn	accuracy	precision	recall	auc	prc
0	1.679126	0.0	91.0	44.0	0.0	0.325926	0.0	0.0	0.0	0.0

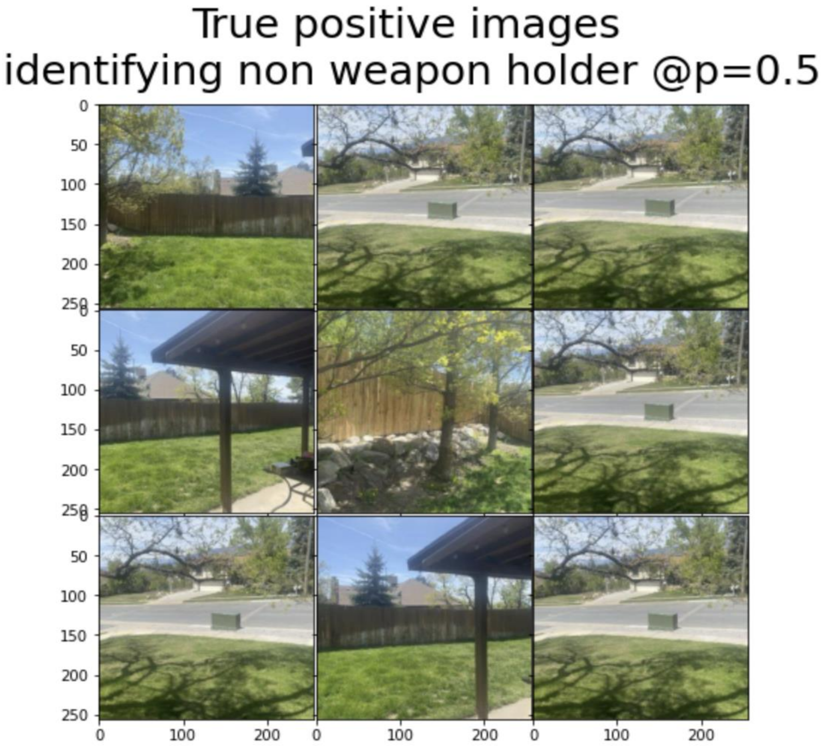
False Positive Adversarial Images

USING AN OUT OF SAMPLE DATASET

- Manually taking photos of a person holding a weapon that the model has not been trained on.
- Replicating weapon situation and placement as closely as possible.
- Performs slightly less than previous in sample modeling.

Accuracy – 0.84
Precision – 0.86
Recall – 0.94





NEXT STEPS - USING LARGER IMAGES AND MORE DATA

- In order to be useful and reliable for any surveillance company, this model would have to perform as close to 100% accuracy.
- Sending an alert that someone is holding a weapon is a very serious topic that needs to be fully developed.
- I am confident that with access to more images and a more stable training dataset that this process could be improved and used in real world applications.
- It's a precarious world and implementations of machine learning like this will be a saving grace.

THANK YOU

MARSHALL@CYRUS1.NET