CLUISE

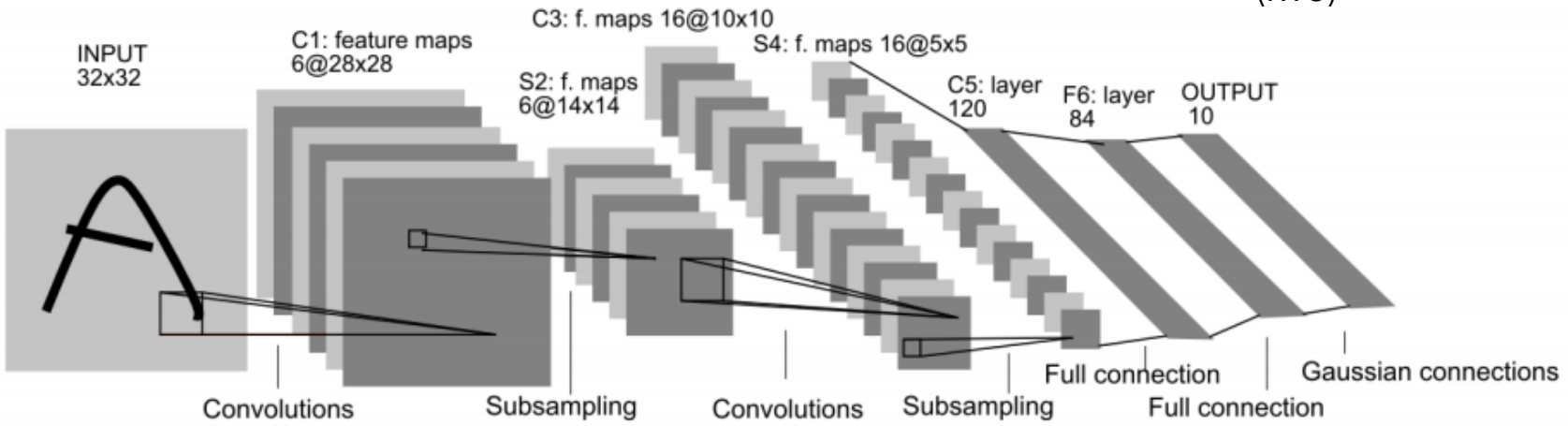
YOUR SMART PERSONAL STYLIST

Categorisation images of clothes with convolutional neural networks (CNN)

Author: Taras Sereda

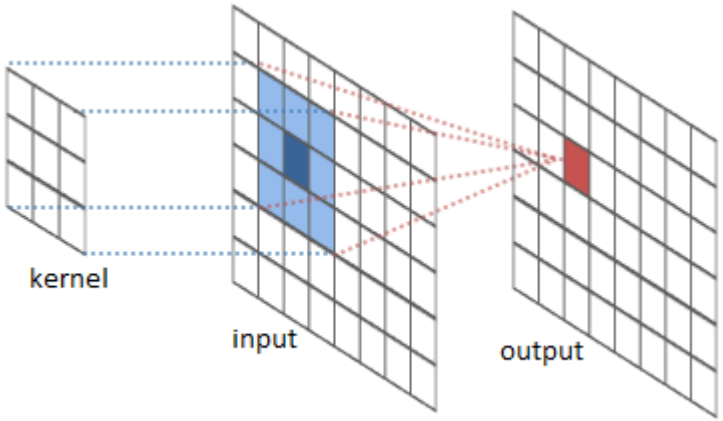
LeNET - 5 (Zip code recognition)

Yan LeCuN
AI lab director at Facebook
(NYU)



[1] http://yann.lecun.com/exdb/lenet/

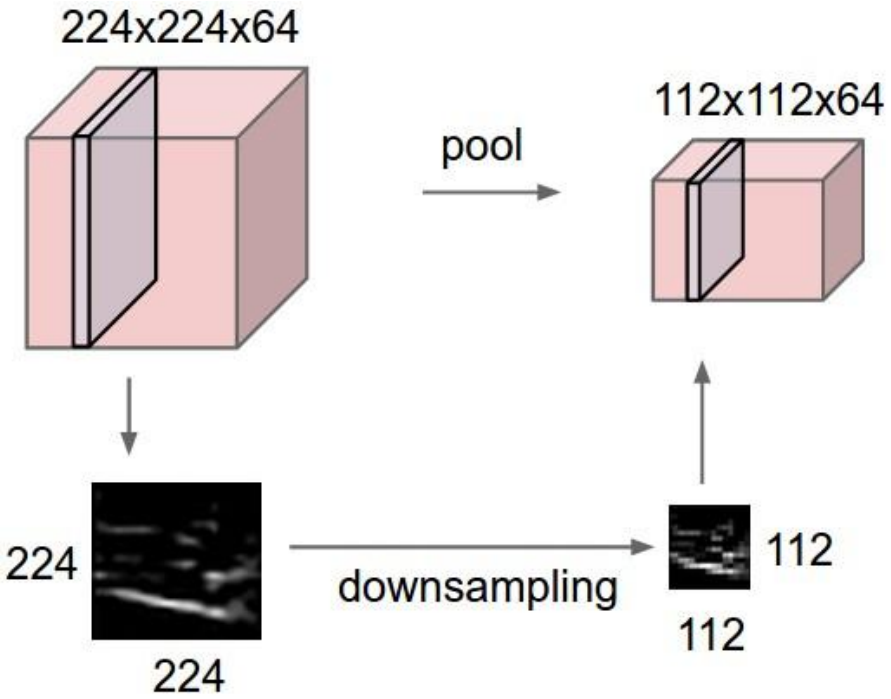
Convolution operation



[2]http://intellabs.github.io/RiverTrail/tutorial/

Dot product between image patch and kernel shifted by bias.

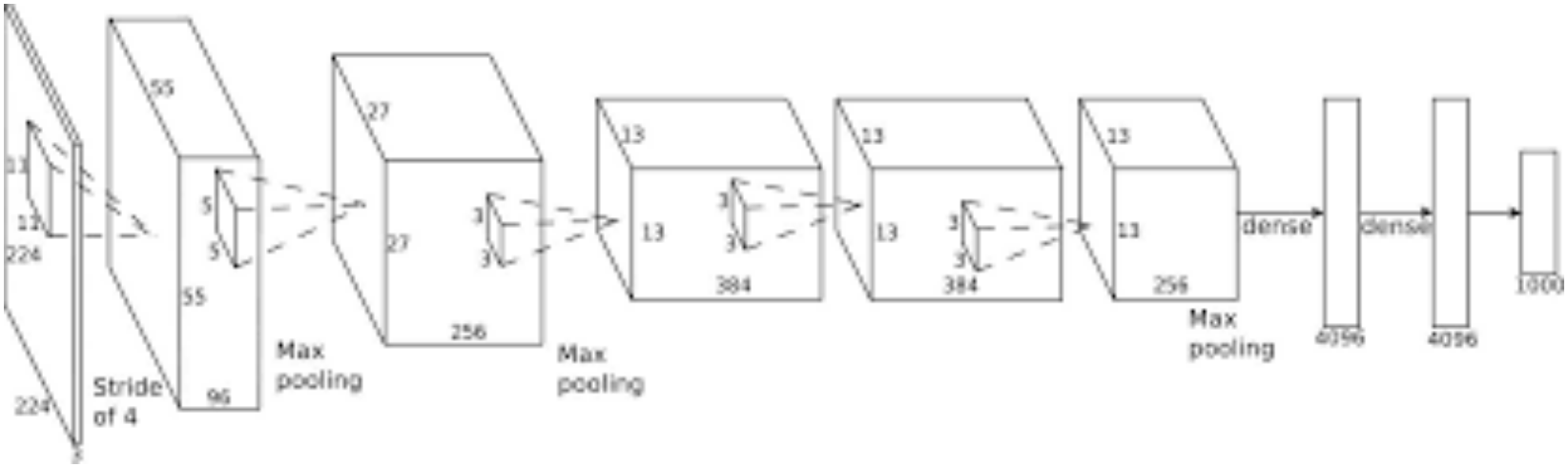
Max Pooling



[3]http://cs231n.github.io/convolutional-networks/

pooling neighboring features, reducing parameters amount.

AlexNet (ImageNet 2012 winning solution)



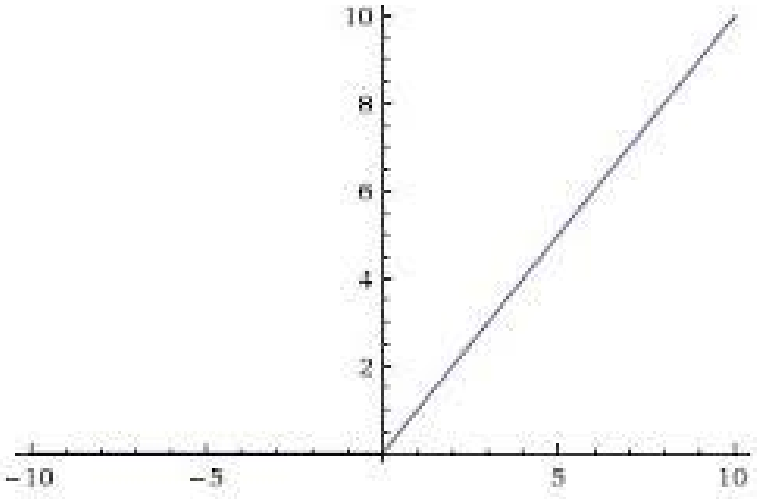
[4] [imagenet-classification-with-deep-convolutional-neural-networks]

[CONV -> POOL] * 2 -> CONV -> CONV -> CONV -> POOL -> FC -> FC -> SOFTMAX

ReLU activation function

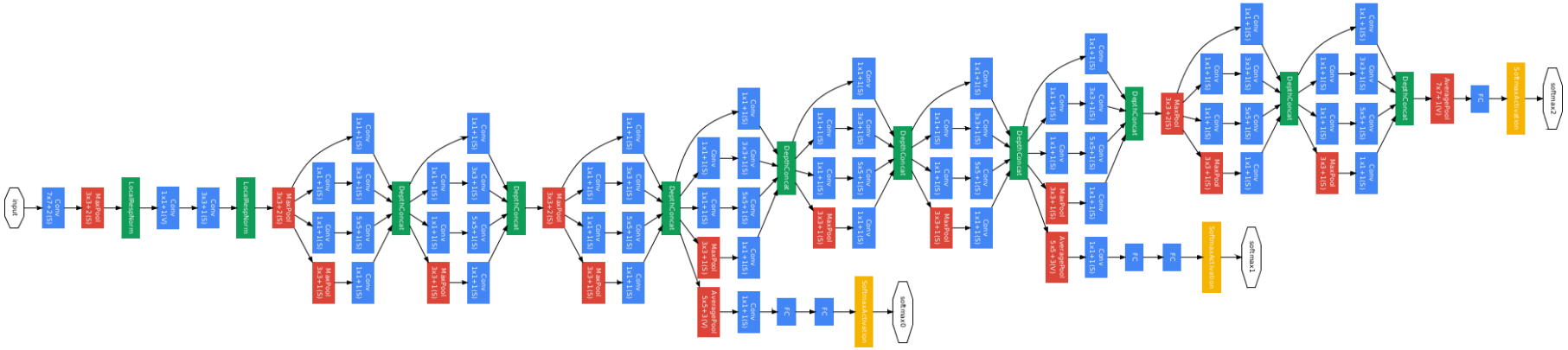
f = max(0,x)

- speeds up training
- more expressive features



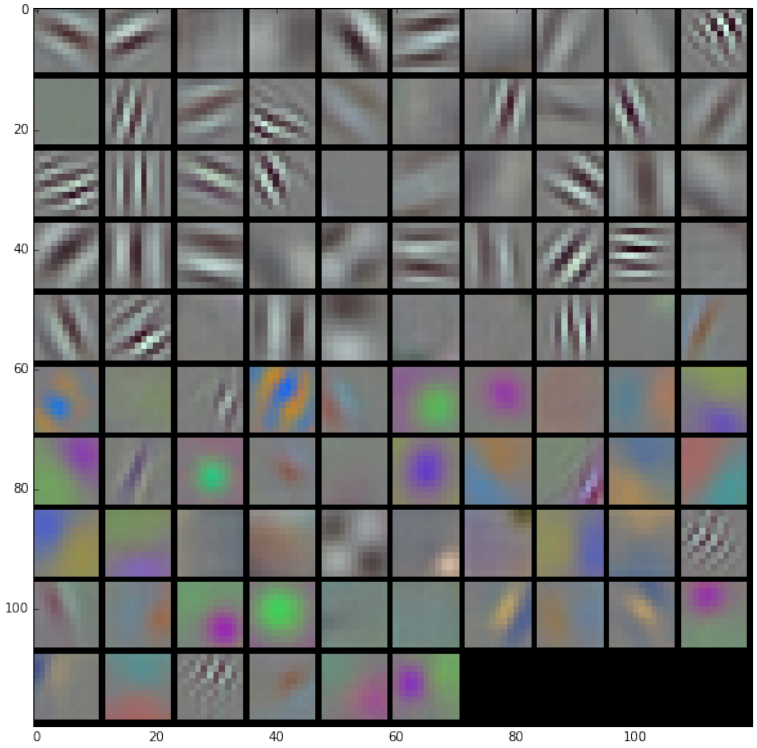
[5]http://cs231n.github.io/neural-networks-1/

GoogleNet (ImageNet 2014 winning solution)



[6] Going deeper with convolution

CONV1 weights



Similar to V1 layer of visual cortex

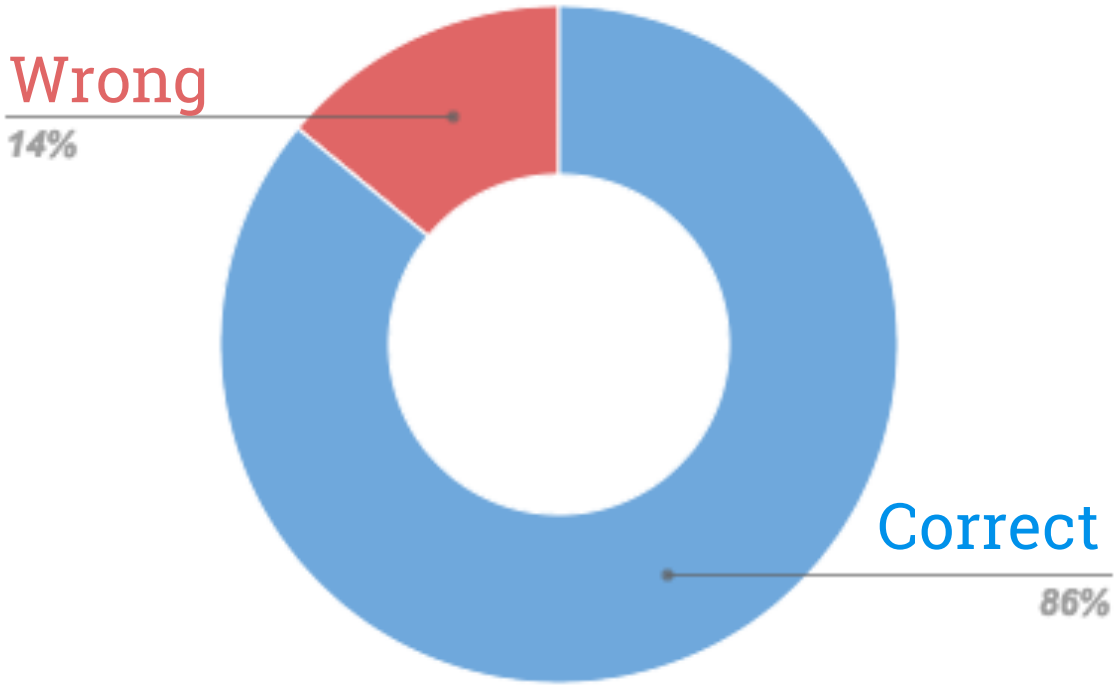
Gathering datasets

- Nearly 800 k of user photos
- ~120 classes in total
- Noize ~30-40%
- manually pick 100 images for each class(human)
- train CNN
- use network for gathering data.

Finetuning for Clothes classification with Caffe

- Small learning rate: 0.001
- Batch_size: 256
- Classes: 56
- Dataset: 20k
- time: 20 hours on K520 GPU
- Amazon: g2.2xlarge

Finetuning Results



1,8k Test set accuracy

Online data augmentation #0

- Random crops
- Random rotations [-50;50]
- For generalization and rotation invariance.

Online data augmentation #1

-90 degree
”accessory.glasses":
"0.78519”

original image
"accessory.glasses”:
“1.0000”

-60 degree
"accessory.glasses":
"0.90990"

Online data augmentation #2

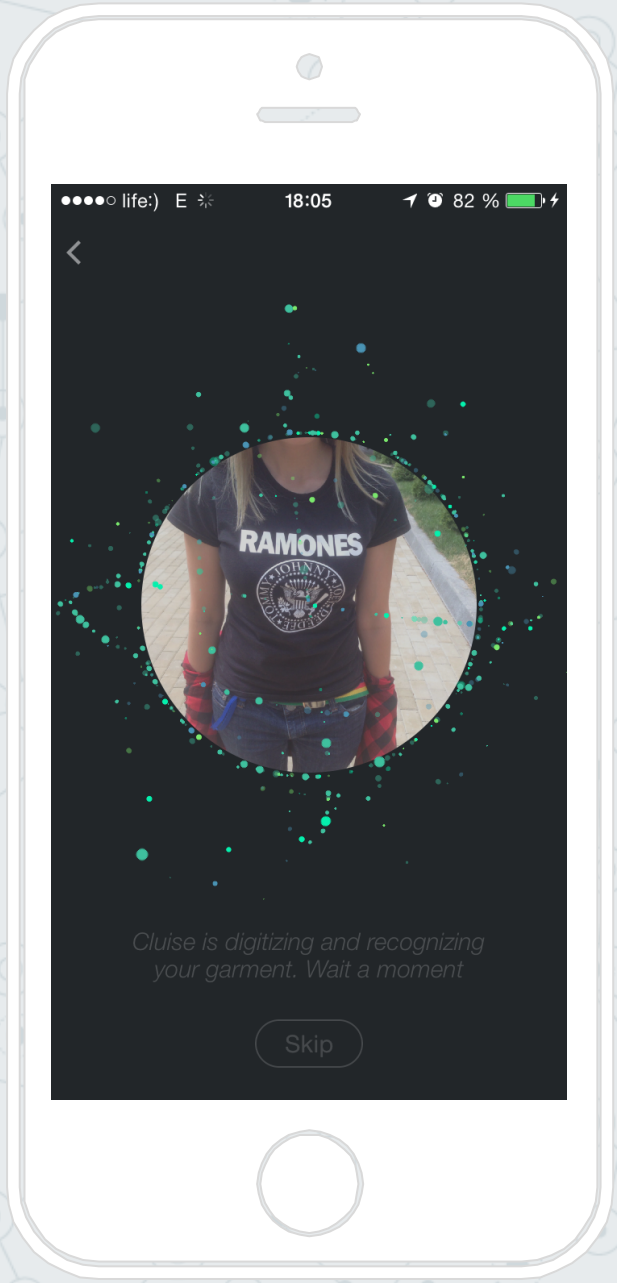
-30 degree
"feet.sport.low trainers": "0.95052"

original image "feet. sport.low trainers": "1.00000"

90 degree
"feet.sport.low trainers": "0.38777",
"feet.sport.high trainers": "0.27408",

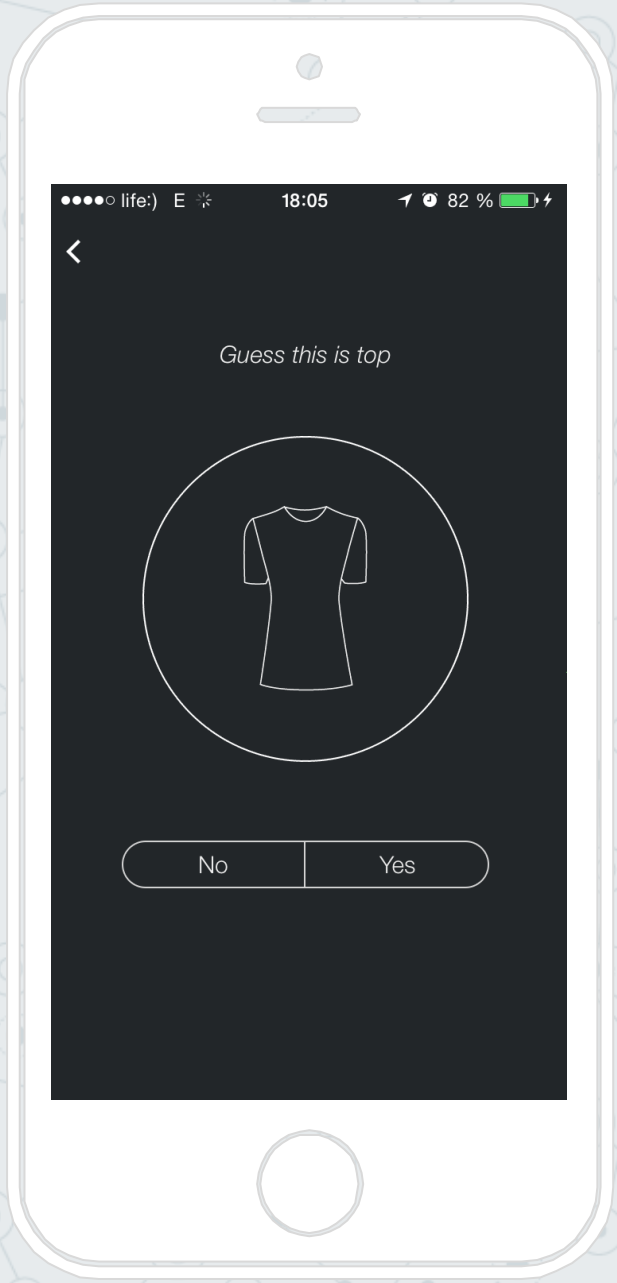
Client side

Garment adding process.

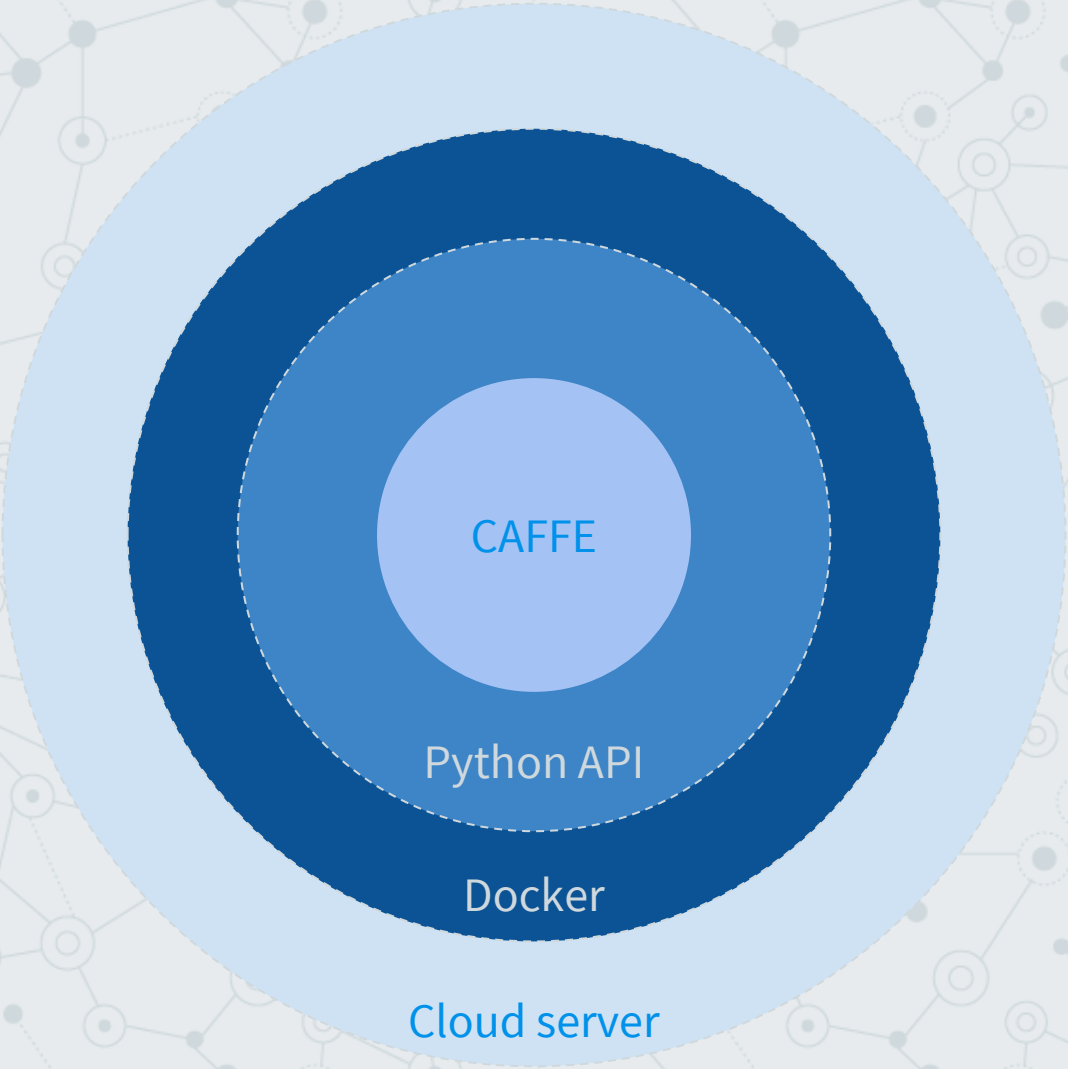


Client side

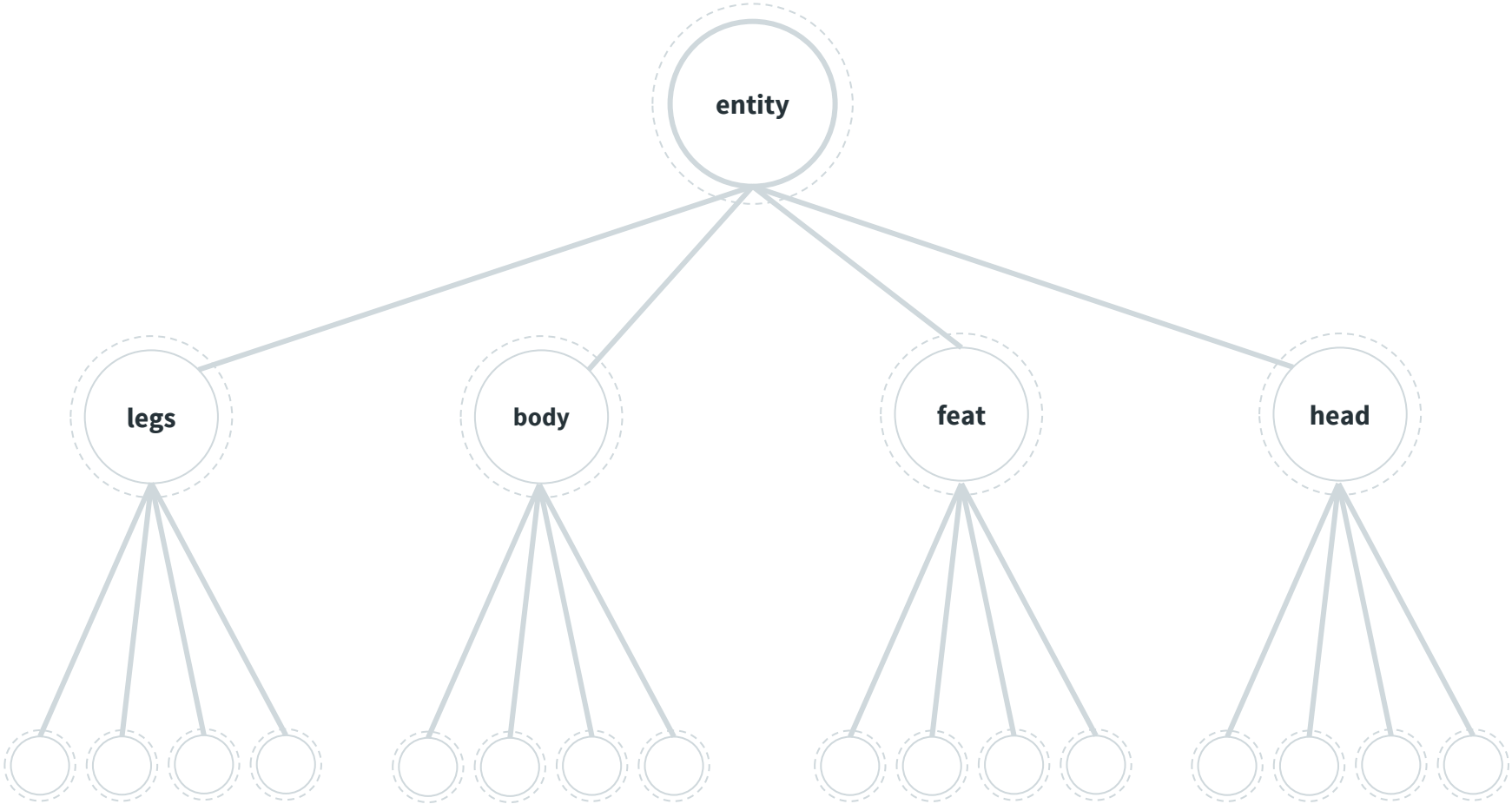
Response as text + icone.



Deployment



Classes Hierarchy



DARTS (hadging your bets)

Solwing accaracy specificity tradeof

$$\begin{aligned} \underset{f}{\text{maximize}} \quad & R(f) \\ \text{subject to} \quad & \Phi(x) \geq 1 - \epsilon \end{aligned}$$

[7] DARTS, hedging your bets

$\Phi(x)$ accuracy

$R(f)$ reward

DARTS Benefits

- get still correct predicts but less specific
- body.knitwear.cardigan is also body.knitwear
- needs only posterior probabilities on leaf nodes.

Future plans

- add more classes,move to 120 in total.
- build ansambel of specialized models
- add image segmentation
- apply attention concepts for image captioning
- move towards online learning

Thanks!

Any questions?

taras.y.sereda@gmail.com