CS230 Hands-on session 2: "Data acquisition and labelling"
Kian Katanforoosh, Andrew Ng
Spring 2018

Part I - "Help the zoo find its iguanas!"

Question 1 (Planning of the work):
You are working with a zoo who asks you to build an algorithm detecting iguanas in real-time. You know that the YOLOv2 algorithm (demoed during the first CS230 lecture) is a great algorithm for real-time object detection. What are the steps to quickly get a working model for the zoo?

Question 2 (Identifying the format of the inherent dataset):
You've identified a repository that implements YOLOv2 in Tensorflow. It is called "darkflow", and the dataset it uses is called "Pascal VOC2012". If you want to use darkflow, what should be the format of your dataset (images and labels)?

Question 3 (You collect the following images):
You went to the zoo and collected the following images:

After asking some object detection experts and searching online, you find a labelling tool called "labelme" that can help you label your images in the format of the "Pascal VOC" dataset. Search labelme on google and find the website, and sign up to access the tool. (Ask the TA for help if you run into issues.)

Take 3 minutes to label the 4 images (all students should do this silently and independently, assume that you are a mechanical turk). Once you're done display simultaneously the 4 labelled images (with the bounding boxes) on your computer screen. Tell the TA when you are ready. PS: If you have time at home, we would advise you to install a labelling tool called "labelImg" (github) and play with it as well.

Question 5:
In groups of 3, discuss and define a common labelling methodology. What did you come up with?

Part II - "NatuRap Language Processing"

Question 1:
You are now working on a natural language processing task. Your goal is to recognise Named Entities in a rap music text such as:

Hey-yo Joe, check my flow!
It's not slow,
It goes faster than YOLO,
Ends up beating Alpha Go.
It's backpropping your network,
Minimizing your net worth.

You find the GitHub repository "sequence_tagging" that does Named Entity Recognition. Take 3min to (individually and independently) label the above text manually so that it matches the github repository. How many classes this problem has? What's the format of the data/annotation?

Question 2:
Assuming you have a huge corpus of rap lyrics, how could you label your data more efficiently? After answering individually, peer-up and discuss your answers in groups of 3.

Question 3:
How would you collect a large dataset to train an algorithm to detect Named Entities in Rap lyrics?

Part III - "Speech RapCognition"

Question 1:
You would like to build a trigger word detection algorithm that detects the word "Hey-yo" in a 10 second audio clip. How would you build a dataset for this task?