Feb13th 2020.

1. Check the result from running model 1a DeepPATH.

Output files:

* nohup.out
* out\_All\_Stats.txt
* out\_filename\_Stats.txt
* precision\_at\_1.txt

Q1: What is precisions in <precision\_at\_1.txt>? How is this value calculated?

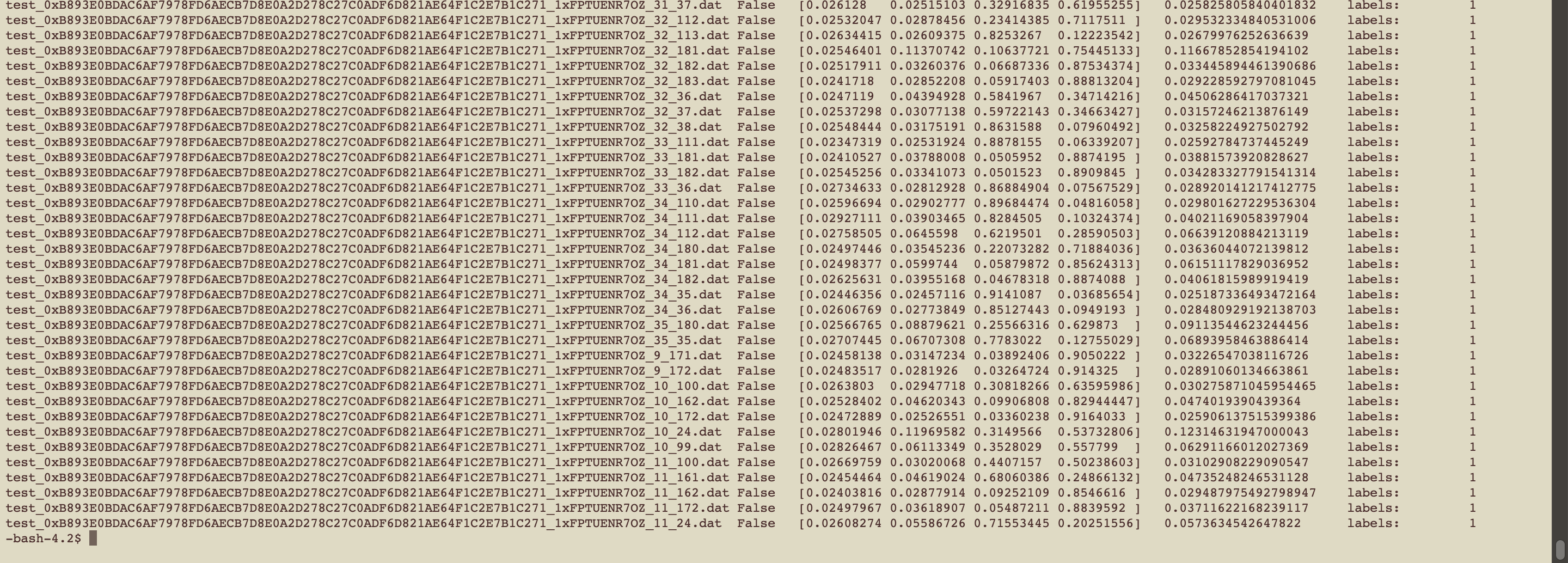
Q2: Labels should be reconsidered: Normal/LUAD/LUAD,=> this is a typo, the author fixed it.

Q3: There are 3 classes, but there are 4 values. Why? And how is this calculated?

Q4: What is the value after [y1, y2, y3, y4]? How is this value calculated?

* out\_filename\_Stats.txt:

<True/False classification> [<output probabilities (with **1st one being the inception's background class**>] <**corrected output probability for the true label** - adjusted to ignore the background class> labels.

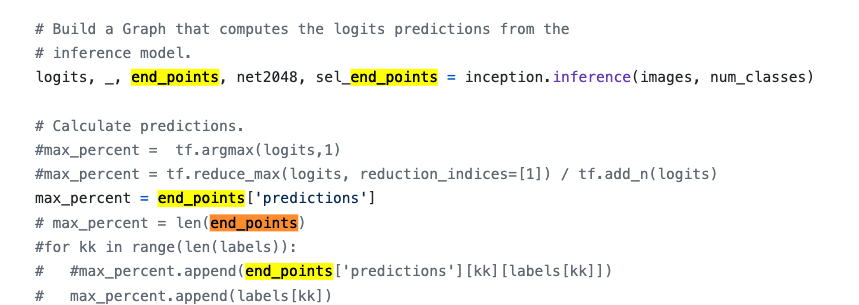


Q5: According to the DeepPATH documentation: <https://github.com/ncoudray/DeepPATH/tree/master/DeepPATH_code#2---run-the-classification-on-the-test-images>

The output file documented at this step is not matching with our output.

* Missing output folder node2048/.
* precision\_at\_1.txt is not mentioning in the documentation. What is this?

**max\_percent:**



**max\_percent** is calculate from the **inception.inference()** function.

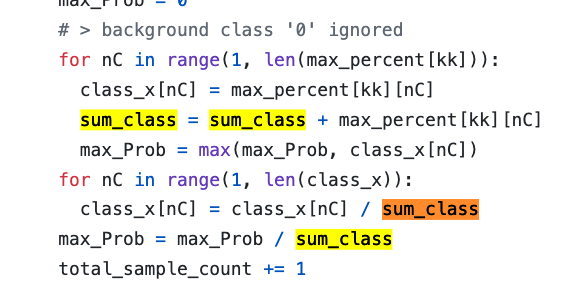
Inception.inferece() is calling inception\_v3() which is the entire architecture of the CNN, and returning the endpoint[‘predictions’], I think it is the results come out of last layer of CNN.

Isn’t the max\_percent is the same for all filenames?

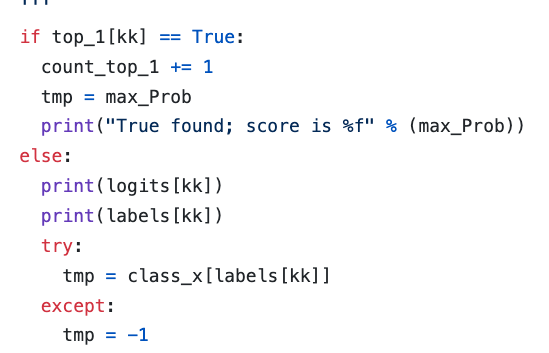
What is **nc** in **class\_x[nc]** ?

Sum\_class

Sess.run()

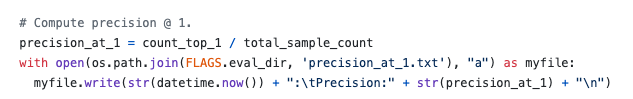
??? Why

If it is FALSE, why print the **logist[kk]**? What is **logists()** do?

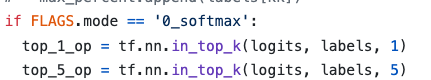




* Above section of code is reason of not generating the output folder 2048/, this means the parameter save == TRUE. But what is the **key** from **endpoint.keys().**



* Precision\_at\_1.txt:
  + What is **count\_top\_1**?



* What is **tf.nn.in\_top\_k()** function?

1. Understand above questions, and go to 3-Analyze the outcome, the goal is to familiarize the whole pipeline. After that we could go back to the beginning to change parameters in pre-processing steps.

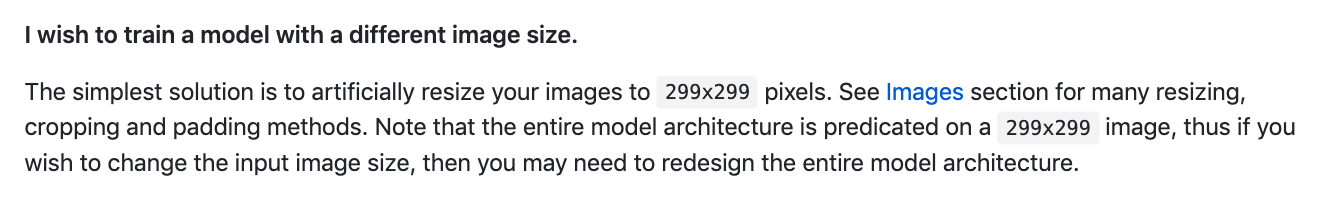
- Based on the documentation, I think the class label on checkpoint documentation is wrong. => issues on GitHub is confirmed with the author and fixed.

- 3 factors:

1. the size of tiles: they used 512, I used 299.

2. background filtering: they used 50, I used 25.

3. magnifications: they used 20x for this check point model, but they indicated that 5x is better.



what is top-1 and top-5 mentioned in Google Inception model?

* <https://stats.stackexchange.com/questions/156471/imagenet-what-is-top-1-and-top-5-error-rate>
* <https://www.quora.com/What-does-the-terms-Top-1-and-Top-5-mean-in-the-context-of-Machine-Learning-research-papers-when-report-empirical-results>
* <https://www.zhihu.com/question/36463511>
* <https://stackoverflow.com/questions/37668902/evaluation-calculate-top-n-accuracy-top-1-and-top-5>
* <https://blog.csdn.net/v1_vivian/article/details/73251187>

maxpool,

AvgPool,

Concat,

Dropout,

Fully connected,

Softmax