



Discussion Forums

Week 9

SUBFORUMS

All

Assignment: Anomaly Detection and Recommender Systems

[← Week 9](#)

Test cases for ex8 - anomaly detection



Tom Mosher · Mentor · Week 9 · 3 years ago · Edited

Here are test cases for the first portion of ex8 - Anomaly Detection:

```
1 Test 1a (Estimate Gaussian Parameters):
2 input:
3 X = sin(magic(4));
4 X = X(:,1:3);
5 [mu sigma2] = estimateGaussian(X)
6 output:
7 mu =
8     -0.3978779    0.3892253   -0.0080072
9 sigma2 =
10     0.27795    0.65844    0.20414
11 -----
12 Test 2a (Select threshold):
13 input:
14 [epsilon F1] = selectThreshold([1 0 0 1 1]', [0.1 0.2 0.3 0.4 0
15     .5]')
16 output:
17 epsilon = 0.40040
18 F1 = 0.57143
19
```

Note: mu and sigma2 may be oriented as either row or column vectors - both are acceptable.

=====



↑ 8 Upvotes

💬 Reply

Follow this discussion

🔒 This thread is closed. You cannot add any more responses.

Earliest

Top

Most Recent

DK Dinesh Kumar · 2 years ago · Edited

Interestingly, i got an issue where / took precedence over *, Though i believed * must happen first. I was getting correct output for TestCase but for exercise. It worked. when overruled the precedence of the operators by using parentheses.

Even the documentation mentions * takes precedence over /.

<https://www.gnu.org/software/octave/doc/v4.0.1/Operator-Precedence.html>

Is it possible to add more test cases?

Also i had a simple calculation issue in F1 score still gave the same results: same Best F1, Best epsilon and 117 anomalies, but was not accepted and assigned grades. I was able to figure out after some time though.

↑ 0 Upvotes

💬 Hide 1 Reply



Paul T Mielke Mentor · 2 years ago

Hi, Dinesh.

I think you are misreading the documentation. The operators "*" and "/" have the same precedence. In that case, the operations are executed from left to right. You can try some experiments and watch what happens:

```
1 3 / 2 * 5
2 3 / (2 * 5)
```

↑ 0 Upvotes

DB Dmitry Bennett · 2 years ago

In case someone struggles with Variance vectorization (needed for "Estimate Gaussian Parameters" part), use

```
1 (bsxfun(@minus, A, B))
```



to get matrix of differences between the X's and the mean - that's how 've done it anyway.



All programming assignments finished - hooray:)

Huge mega thanks to Andrew Ng and Tom Mosher for this brilliant, brilliant course!

↑ 7 Upvotes

💬 Reply



José López Liévanos · 3 years ago



This is just a suggestion for all of you, try not to use the classperf function and instead do it like in the implementation note says.

↑ 1 Upvote

💬 Hide 1 Reply



Tom Mosher Mentor · 3 years ago



No problem, it doesn't exist in Octave.

↑ 2 Upvotes

FZ

Fanhui Zeng · 3 years ago



Hey Tom.

In the second case, when epsilon is 0.1004, F1 equals to 0.75, much larger than 0.57143. So maybe there are something wrong in this case.

↑ 0 Upvotes

💬 Hide 2 Replies



Tom Mosher Mentor · 3 years ago



Your F1 value is incorrect, see my previous reply (from 4 minutes ago).

↑ 0 Upvotes



Tommy Wong · 2 years ago



You may typo the recall formula to $tp/(fp+fn)$, that's my case. And thanks, Tom.

↑ 3 Upvotes

FZ

Fanhui Zeng · 3 years ago



Hey, Tom,

I think the second test case have something wrong. When epsilon is 0.1004, the prediction vector is $[1 \ 0 \ 0 \ 0 \ 0]$, so the $tp = 1$, $fp = 0$ and $fn = 2$. Then $prec = 1$, $rec = 3/5$. So $F1 = 0.75$, much larger than 0.57143.

↑ 0 Upvotes Hide 1 Reply



Tom Mosher · Mentor · 3 years ago

I believe the test case is correct.

When epsilon is 0.1004, $tp = 1$, $fp = 0$, and $fn = 2$.

↑ 2 Upvotes



Steve Thollar · 3 years ago

I am confused. In the output of your test case for `estimateGaussian(X)` μ and σ^2 are 1×3 , but in the `estimateGaussian` function these appear that they should be 3×1 . In particular in initializing μ and σ^2 , it has the commands $\mu = \text{zeros}(n,1)$ and $\sigma^2 = \text{zeros}(n,1)$. When I run your test case, my output is the transpose of your output. Am I missing something???

↑ 0 Upvotes Hide 3 Replies



Tom Mosher · Mentor · 3 years ago

I understand your confusion. It confuses me too.

The pre-defined return value for μ in the function template is in fact $(n \times 1)$. But when you compute the μ and σ values, they are naturally oriented with the columns of X , so a $(1 \times n)$ orientation is actually more efficient - they will have to be used as a row vector anyway, so transposing them before returning the value doesn't really add much value.

The functions that use the return values from the `estimateGaussian()` function - `multivariateGaussian()` and `visualizeFit()` - have been written to accept μ and σ in either orientation. And the submit grader ignores the orientation of the return values in all cases.

So it turns out that either orientation works correctly, and you may implement it either way.

↑ 1 Upvote



Tom Mosher · Mentor · 3 years ago

I will update the test case to make this more clear. Thanks for pointing out the discrepancy.



↑ 0 Upvotes

coursera



Steve Thollar · 3 years ago



No probs. Thanks for the quick response.

↑ 0 Upvotes



Fred Leve · 3 years ago



I am getting an error

"warning: division by zero" and the error

"!! Submission failed: unexpected error: urlread: HTTP response code said error

!! Please try again later."

when I try to submit. I checked all of the gradients, costs, and values that I should have obtained and they were exact. I do not understand why I am getting this error. I also changed Y to Ynorm in line 200 of ex8_cofi.m. I am not sure why it is not working.

↑ 0 Upvotes

Hide 1 Reply



Tom Mosher Mentor · 3 years ago · Edited



The "urlread" error means the submit script is not able to get a connection with the Coursera server that accepts your exercise results.

Typically students see this problem earlier in the course. It is very common when doing ex1, then students know how to handle it.

If the first time you are having this problem is now in ex8, then perhaps something has changed in your computer setup.

In any case, the work-arounds for this are given in the thread "Mentor tips for submitting your work" in the General Discussion forum area.

↑ 2 Upvotes



Charles Snyder · 3 years ago · Edited



Has anyone run into the mistake of calculating the correct epsilon, but getting the wrong F1 value? If so, could you share any hints on where your calculation of F1 was incorrect? For example, in this test case, I get epsilon = 0.40040, but my F1 value is 0.72727. (I get to correct epsilon of 8.99e-05 for ex8 also, making this a bit tricky to find the error.) I am calculating F1 as $2 * (\text{prec} * \text{rec}) / (\text{prec} + \text{rec})$;



↑ 0 Upvotes

Hide 4 Replies

coursera



Tom Mosher · Mentor · 3 years ago · Edited



The only trick here is getting the logic correct combinations (of the predictions and yval) for the fp, tp, and fn values. After that, the rest is just algebra.

↑ 1 Upvote



Charles Snyder · 3 years ago



Ah, thank you. Yes, I read tp "as the number of true positives," but failed to add the extra logic "...and our algorithm correctly classified it as an anomaly." Good to go now!

↑ 8 Upvotes



Suya Liu · 3 years ago



lol...thanx Charles, I had the exact same problem... that'll teach me to finish reading all the instructions!

↑ 1 Upvote

VS

Vivek Sasikumar · 2 years ago



Exact same error. Thanks guys.

↑ 0 Upvotes

LK

Larry Kline · 3 years ago



I get a bunch of divide by zero warnings, however my code for Select Threshold passes. Should I put a 'continue' statement in the for loop if the denominator of an expression is zero?

↑ 0 Upvotes

Hide 4 Replies



Tom Mosher · Mentor · 3 years ago



Warnings for divide by zero are not usually a problem. Octave/MATLAB handles it gracefully.

↑ 0 Upvotes



Mwombeki F. Fabian · 3 years ago



Tom,



I have written a for loop though `i=1:size(pval,1)` to return `yval(i)` as my prediction if `(p<epsilon)`. For some reason I keep getting zeros. Any insights?

Coursera



↑ 0 Upvotes



Tom Mosher · Mentor · 3 years ago



The Note in `selectThreshold.m` says to use this:

```
1 predictions = (pval < epsilon)
```

No for-loop is required.

↑ 1 Upvote



Mwombeki F. Fabian · 3 years ago



Thanks. Made it!

It's simpler than I had approached it.

↑ 0 Upvotes

RZ

Reagan Zogby · 3 years ago



Working the second part of the Anomaly detection problem It involves calculating F1. This in turn I believe means knowing now how many defective units are in the sample space. Is that correct, and if so, where should we be looking to get that information? I may be missing something obvious. Thanks.

↑ 0 Upvotes

💬 Hide 2 Replies



Tom Mosher · Mentor · 3 years ago



See `ex8.pdf` on page 6, especially the text box on the bottom of the page.

↑ 2 Upvotes

RZ

Reagan Zogby · 3 years ago



Thxs. Appreciate it.

↑ 0 Upvotes

