Questions

Menu

- Questions
- Branching
- Custom question numbers
- Set page breaks
- Remove page breaks
- Preview survey
- Reports and data
- Survey home

Recent items

- mehilshah
- Group - 2E
- Group - 2C
- Group - 1E
- Group - 1C
- Group - 4E
- Group - 3E
- raisedal
- main folder

Guide

The following steps are recommended for your survey:

- Create questions
- Add conditional branching
- Customize look and feel
- Set privacy and behavior
- Translate survey
- Publish survey
- Analyze collected data

Group - 3C

Section 1 Edit

New text

Introduction

Hello there!
Welcome to this survey! We are a group of researchers from Dalhousie University, Canada. Recently, we conducted an empirical study involving 85 reproducible bugs from Stack Overflow posts. Our aim was to understand two main aspects: (1) the edit actions that can be employed to complete code snippets for bug reproduction and (2) the information that enhances the reproducibility of bug reports. Our investigation has yielded several interesting findings, and we are seeking your feedback on them.

We reproduced 85 bugs and discovered they could be reproduced using 10 edit actions. To enhance their reproducibility, 5 main information categories need to be present. The edit actions and information categories are described below.

Edit Actions
Input Data Generation: Generating input data which simulates the data used for training the model.
Neural Network Construction: Reconstructing or modifying the neural network based on the information provided
Hyperparameter Initialization: Initializing the hyperparameters for training, such as batch size and number of epochs
Import Addition and Dependency Resolution: Determining the dependencies in the code snippet and adding the missing imports.
Logging: Adding appropriate logging statements to capture relevant information during reproduction
Obsolete Parameter Removal: Removing outdated parameters or functions to match the parameters of the latest library versions
Compiler Error Resolution: Debugging and resolving compiler errors that arise due to the errors in the provided code snippet.
Dataset Procurement: Acquiring the datasets and using them to train the model
Downloading Models & Tokenizers: Fetching pre-trained models and tokenizers from external sources.
Version Migration: Updating the code to adapt the changes introduced in newer library or framework versions.

Information Categories
Data: Shape of the input data, type of data, data distribution.
Model: Neural network architecture, number of layers, neurons, activation function for layers.
Hyperparameters: Batch size, epochs, optimizers, loss function.
Code Snippet: Training code snippet, evaluation script, data preprocessing, and transformation operations.
Logs: Compiler error logs, training error logs

[Edit | Delete]

---------- page break ----------

Section 2 Edit | Delete

New text

Demographics

New text | New question | New question from library / other surveys

Question 1 Edit | Add to library | Delete

1. What is your relevant work experience with deep learning?

- ◯ <1 Year
- ◯ 1-5 Years
- ◯ 5-10 Years
- ◯ >10 Years

Split section | New text | New question | New question from library / other surveys

Question 2 Edit | Add to library | Delete

2. What is your relevant experience with deep learning bug fixing?

- ◯ <1 Year
- ◯ 1-5 Years
- ◯ 5-10 Years
- ◯ >10 Years

Split section | New text | New question | New question from library / other surveys

Question 3 Edit | Add to library | Delete

3. What is your current occupation?

- ◯ Software Practitioner (Software Engineer, Deep Learning Engineer, Machine Learning Engineer etc.)
- ◯ Researcher (Masters/Doctoral Student, PostDoc, Faculty)

Split section | New text | New question | New question from library / other surveys

Question 4 Edit | Add to library | Delete

4. What are the deep learning frameworks you have worked with?

- [ ] Tensorflow
- [ ] PyTorch
- [ ] Keras
- [ ] Other

Split section | New text | New question | New question from library / other surveys

Question 5 Edit | Add to library | Delete

5. What challenges are associated with reproducing deep learning bugs in your day-to-day activities?

New text

---------- page break ----------

Section 3 Edit | Delete

Bug #1

Given the issue description, and the code snippet. Please reproduce the bug, select the most appropriate edit operations and critical information needed to reproduce this bug.

To help the reproduction process, we have provided the sample edit operations here.

Original Issue Report: https://stackoverflow.com/questions/69549126/tensorflow-typeerror-cannot-convert-1e-12-to-eagertensor-of-dtype-int32

Description:

I have a multiclass classification machine learning application for which I want to calculate the f1 score using tensorflow. The predicted and actual values are stored in pandas dataframes y_pred and y_act respectively. Both are populated with 1's and 0's. So I do something like this (Code Snippet provided below):

However I get the following error

TypeError: Cannot convert 1e-12 to EagerTensor of dtype int32

There must be something with the type casting from pandas to tensorflow which is throwing the error. I have tried a series of mitigations to no avail.

I tried converting the numpy arrays to tensors like so: pred_tf = tf.convert_to_tensor(pred_numpy, numpy.int32)

I tried ensuring the pandas dataframe has no 1e-12 instances with: y_pred = y_pred.replace(1e-12, 0)

I tried converting to numpy without the numpy.int32 option.

However I still get the same error. Any tips for converting from pandas to tensors successfully without getting this error?

Code Snippet: You can use this Colab notebook as the base notebook to start the reproduction process: https://colab.research.google.com/drive/14yuLcsJ6vg0ECahKeUcQ1BuORi3kkJsZ?usp=sharing

[Edit | Delete]

New text | New question | New question from library / other surveys

Question 6 Edit | Add to library | Delete

6. What are the edit operations that could be used to reproduce this bug?

- [ ] Input Data Generation
- [ ] Hyperparameter Initialization
- [ ] Logging
- [ ] Compiler Error Resolution
- [ ] Downloading Models and Tokenizers
- [ ] Neural Network Construction
- [ ] Import Addition and Dependency Resolution
- [ ] Obsolete Parameter Removal
- [ ] Dataset Procurement
- [ ] Version Migration

Split section | New text | New question | New question from library / other surveys

Question 7 Edit | Add to library | Delete

7. Why do you think these edit operations could prove useful in reproducing the bug?

Split section | New text | New question | New question from library / other surveys

Question 8 Edit | Add to library | Delete

8. What are the critical information components that could help the reproducibility of this bug?

- [ ] Data
- [ ] Hyperparameters
- [ ] Model
- [ ] Code Snippet
- [ ] Logs

Split section | New text | New question | New question from library / other surveys

Question 9 Edit | Add to library | Delete

9. How do you think the selected critical information could be useful in reproducing the bug?

Split section | New text | New question | New question from library / other surveys

Question 10 Edit | Add to library | Delete

10. Did you implement any additional operations or actions beyond those suggested by us? Please let us know your thoughts.

New text

--------- page break ---------

Section 4 Edit | Delete

Bug #2

Given the issue description, and the code snippet. Please reproduce the bug, select the most appropriate edit operations and critical information needed to reproduce this bug.

To help the reproduction process, we have provided the sample edit operations here.

Original Issue Report: https://stackoverflow.com/questions/54417736/pytorch-runtime-error-invalid-argument-0-sizes-of-tensors-must-match-except-i

Description:

I have a PyTorch model and I'm trying to test it by performing a forward pass. The model is defined below.

If I try:

```
modl = ResUnet((1,320,320), 1)
x = torch.rand(1, 1, 320, 320)
modl(x)
```

It throws an error

```
---------------------------------------------------------------------------
RuntimeError                              Traceback (most recent call last)
<ipython-input-46-4ddc821c365b> in <module>
----> 1 modl(x)

~/.conda/envs/torch0.4/lib/python3.6/site-packages/torch/nn/modules/module.py in __call__(self, *input, **kwargs)
    475             result = self._slow_forward(*input, **kwargs)
    476         else:
```

```
--> 477             result = self.forward(*input, **kwargs)
    478         for hook in self._forward_hooks.values():
    479             hook_result = hook(self, input, result)

<ipython-input-36-f9eeefa3c0b8> in forward(self, x)
    221         de2_ = self.d2(de1)
    222         #print de2_.size()
--> 223         de2 = torch.cat([en6add,de2_],1)
    224         #print de2.size()
    225
```

RuntimeError: invalid argument 0: Sizes of tensors must match except in dimension 1. Got 5 and 4 in dimension 2 at /opt/conda/conda-bld/pytorch_1535491974311/work/aten/src/TH/generic/THTensorMath.cpp:3616
I figure the problem is caused by the input size not being a power of 2 but I am not sure how to rectify it for the given input dimenstions (320, 320).

Code Snippet: You can use this Colab notebook as the base notebook to start the reproduction process: https://colab.research.google.com/drive/1hkA6aDFEN3CG7w4qPVijjn1W3IfwEeqJ#scrollTo=J_guUFDERox4

[Edit | Delete]

New question | New question from library / other surveys

Question 11 Edit | Add to library | Delete

11. What are the edit operations that could be used to reproduce this bug?

- [ ] Input Data Generation
- [ ] Hyperparameter Initialization
- [ ] Logging
- [ ] Compiler Error Resolution
- [ ] Downloading Models and Tokenizers
- [ ] Neural Network Construction
- [ ] Import Addition and Dependency Resolution
- [ ] Obsolete Parameter Removal
- [ ] Dataset Procurement
- [ ] Version Migration

Split section | New text | New question | New question from library / other surveys

Question 12 Edit | Add to library | Delete

12. Why do you think these edit operations could prove useful in reproducing the bug?

Split section | New text | New question | New question from library / other surveys

Question 13 Edit | Add to library | Delete

13. What are the critical information components that could help the reproducibility of this bug?

- [ ] Data
- [ ] Hyperparameters
- [ ] Model
- [ ] Code Snippet
- [ ] Logs

Split section | New text | New question | New question from library / other surveys

Question 14 Edit | Add to library | Delete

14. How do you think the selected critical information could be useful in reproducing the bug?

Split section | New text | New question | New question from library / other surveys

Question 15 Edit | Add to library | Delete

15. Did you implement any additional operations or actions beyond those suggested by us? Please let us know your thoughts.

New text

---------- page break ----------

Section 5 Edit | Delete

New text

Unique ID Generation

New text | New question | New question from library / other surveys

Question 16 Edit | Add to library | Delete

16. Please use this secure link, and enter the Unique ID generated in the following textbox.

If you want to withdraw from the survey, email us with this Unique ID at shahmehil@dal.ca, and we will delete your response.

Unique ID

New text | New question | New question from library / other surveys

---------- page break ----------