Questions

Menu

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

Recent items

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

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 - 4E

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Obsolete Parameter Removal: Removing outdated parameters or functions to match the parameters of the latest library versions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/69137834/keras-valueerror-dimensions-must-be-equal-but-are-2-and-32-for-node-equal

Description:

I was trying to train a simple Keras network for classification when I faced the following error. I know there is something wrong with my inputs but I couldn't figure out how to fix it. Here is my code (given below)

Dataset Shape:
x_train : float32 0.0 1.0 (2444, 64, 64, 1)
y_train : float32 0.0 1.0 (2444, 2)
x_test : float32 0.0 1.0 (9123, 64, 64, 1)
y_test : float32 0.0 1.0 (9123, 2)

Error:
ValueError: Dimensions must be equal, but are 2 and 32 for '{{node Equal}} = Equal[T=DT_FLOAT, incompatible_shape_error=true] (IteratorGetNext:1, Cast_1)' with input shapes: [?,2], [?,32,32].

Code Snippet: You can use this Colab notebook as the base notebook to start the reproduction process: https://colab.research.google.com/drive/1eUZNzpmNhadBxC2hW1bZSOgVv8tOXxXG?usp=sharing

Hints:

1. Data, Logs, and Code Snippet can be useful information for reproducing the bug.
2. Import Addition, Input Data Generation, Hyperparameter Initialization, Dataset Procurement, and Logging can be useful edit actions for reproducing the bug.
3. Focus on the statement: "Dataset Shape: x_train : float32 0.0 1.0 (2444, 64, 64, 1), y_train : float32 0.0 1.0 (2444, 2), x_test : float32 0.0 1.0 (9123, 64, 64, 1), y_test : float32 0.0 1.0 (9123, 2)".

[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/64576751/neural-network-typeerror-unsupported-operand-types-for-dense-and-str

Description:

I am trying to use a neural network to predict the price of houses. Here is what the top of the dataset looks like:

```
   Price  Beds   SqFt   Built  Garage  FullBaths  HalfBaths  LotSqFt
   485000 3      2336   2004   2       2.0        1.0        2178.0
```

```
430000  4      2106   2005   2     2.0        1.0       2178.0
445000  3      1410   1999   1     2.0        0.0       3049.0
```

...
I am using the ReLU activation function. When I try to evaluate my model on my test data, I get this TypeError: unsupported operand type(s) for +=: 'Dense' and 'str'.

I looked at the types of the columns from my original dataframe, and everything looks fine.

```
print(df.dtypes)
## Output
#Price         int64
#Beds          int64
#SqFt          int64
#Built         int64
#Garage        int64
#FullBaths     float64
#HalfBaths     float64
#LotSqFt       float64
#dtype: object
```

I'm not sure if I am messing something up in my neural network to cause this error. Any help is appreciated

Code Snippet: You can use this notebook as the base notebook to start the reproduction process: https://colab.research.google.com/drive/127nDAw8gnh9yDpZmHSFDRdDP1Mv7hX6q?usp=sharing

Hints:

1. Logs, Code Snippet and Model can be useful information for reproducing the bug.
2. Input Data Generation, Hyperparameter Initialization, Import Addition, Logging, and Obsolete Parameter Removal can be useful edit actions for reproducing the bug.
3. Focus on the statement: "When I try to evaluate my model on my test data, I get this TypeError: unsupported operand type(s) for +=: 'Dense' and 'str'. print(df.dtypes), ## Output, #Price - int64, #Beds - int64, #SqFt - int64, #Built - int64, #Garage - int64, #FullBaths - float64, #HalfBaths - float64, #LotSqFt - float64, #dtype: object"

[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 ---------