

# **Predicting coding question quality using Stack Overflow ratings**

IMT4904 Master thesis – Spring 2016

Master Applied Computer Science – Web, Mobile, Games

130533 – Knut Lucas Andersen

# Overview

- [Stack Overflow \(SO\)](#)
- [What is a question?](#)
- [Support Vector Machine \(SVM\)](#)
- [Methodology](#)
- [Experiments and Results](#)
- [Summary](#)
- Demo
- [References](#)

# Stack Overflow (SO)

- Released late 2008
- A Question-Answering (QnA) community for programmers
- Model which Stack Exchange is built on
- Uses gamification to reward users for participation
  - Reputation
  - Votes
  - Badges
  - Accepted answer
  - Bounty
- Who are the experts?
- What defines a good Question and Answer on SO?

# How do you undo the last commit?

Add  projects to your  **stackoverflow** profile.

▲ I committed the wrong files to Git.

10493 How can I undo this commit?

▼ [git](#) [git-commit](#) [git-reset](#) [git-revert](#)

★ share edit  
3886

edited Apr 14 at 21:01

community wiki  
40 revs, 25 users 25%  
[Peter Mortensen](#)

105 Starting with Git v1.8.4, all the answers below that use `HEAD` or `head` can now use `@` in place of `HEAD` instead. See [this answer \(last section\)](#) to learn why you can do that. – [user456814](#) Jul 26 '13 at 2:05

55 Amend the commit, and use a [gitignore](#) to stop it happening again. – [Zaz](#) Aug 12 '13 at 13:27

9 Consider installing a [git-extras](#) [github.com/visionmedia/git-extras](#) tool, it have `git undo` task that can take a number of steps to undo as an argument, like: `git undo 3`. – [Dan K.K.](#) Nov 14 '13 at 16:29

5 See also [Revert to previous Git commit](#). – [user456814](#) May 23 '14 at 17:56

8 why are there so many ways of doing it!? its just makes it so confusing... its simple, we accidentally made a git commit to a changed file.. how do you just undo that command? – [duckx](#) Dec 22 '14 at 5:52

[show 6 more comments](#)

45 Answers

active oldest votes

1 2 next

▲ Undo a commit and redo

10393

▼  
✓  

```
$ git commit -m "Something terribly misguided" (1)
$ git reset --soft HEAD~ (2)
<< edit files as necessary >> (3)
$ git add ... (4)
$ git commit -c ORIG_HEAD (5)
```

1. This is what you want to undo

2. This is most often done when you remembered what you just committed is incomplete, or you misspelled your commit message<sup>1</sup>, or both. Leaves working tree as it was before `git commit`.

asked 7 years ago

viewed 3304185 times

active today

## HOT META POSTS

- 228 [Should Stack Overflow be more restrictive about new user registrations?](#)
- 6 [Merge the \[gestures\] into \[gesture\]?](#)
- 12 [Funky way to deal with a duplicate question?](#)
- 95 [The life of answering questions on Stack Overflow](#)

Want a [python](#) job?

### Python Developer

WeTransfer  Amsterdam, Netherlands

[python](#) [airflow](#)

### Senior Data Engineer - Python (m/f)

Delivery Hero  Berlin, Germany

 RELOCATION

[python](#) [git](#)

### Senior Backend Developer (f/m)

resmio  Berlin, Deutschland

€45,000 - €55,000

[python](#) [django](#)

### Software Developer

Optiver  Amsterdam, Netherlands

 RELOCATION  VISA SPONSORSHIP

[python](#) [c++](#)

Linked

43 [Is there a way to rollback my last push to Git?](#)

[1]

Work on work you love. From home.



How can I create a loop that goes back to the beginning stage? My issue is here:

-154



15

```
int main(void)
{
    midi_start();
    program_change(1, 1);
    program_change(2, 1);
    t=400;

    printf("Choose a scale and write the code of it:\n ");
    printf("C:0\n ");
    printf("C#:1\n ");
    printf("D:2\n ");
    printf("D#:3\n ");
    printf("E:4\n ");
    printf("F:5\n ");
    printf("F#:6\n ");
    printf("G:7\n ");
    printf("G#:8\n ");
    printf("A:9\n ");
    printf("A#:10\n ");
    printf("B:11\n ");
    printf("The chosen key is:\n");
    scanf("%d",&kk);
    k = kk + 60;
    if (kk >= 12)
    {
        printf("Go back to the start.");
        return main;
    }
}
```

I want to make it return to the "Please, choose a key" command so that the person using this can correct the mistake.

c

share edit

edited Jan 16 at 15:26

Mark Amery  
19.3k • 10 • 101 • 130

asked Jan 11 '14 at 22:35

user3186187  
19 • 1 • 8

locked by Brad Larson ♦ May 16 at 14:30

This post has been locked while disputes about its content are being resolved. For more info [visit meta](#).

closed as off-topic by πάντα ῥεῖ, rene, Martin James, Bjørn-Roger Kringsjå, Tunaki May 14 at 17:06

This question appears to be off-topic. The users who voted to close gave this specific reason:

- "This question was caused by a problem that can no longer be reproduced or a simple typographical error. While similar questions may be on-topic here, this one was resolved in a manner unlikely to help future readers. This can often be avoided by identifying and closely inspecting the [shortest program necessary to reproduce the problem](#) before posting." – πάντα ῥεῖ, rene, Martin James, Bjørn-Roger Kringsjå, Tunaki

If this question can be reworded to fit the rules in the [help center](#), please [edit the question](#).

# Question Badges

● Altruist	First bounty you manually award on another person's question	6.3k awarded
● Benefactor	First bounty you manually award on your own question	30.3k awarded
● Curious	Ask a well-received question on 5 separate days, and maintain a positive question record	213.8k awarded
● Inquisitive	Ask a well-received question on 30 separate days, and maintain a positive question record	21.4k awarded
● Socratic	Ask a well-received question on 100 separate days, and maintain a positive question record	2.4k awarded
● Favorite Question	Question favorited by 25 users	28.2k awarded
● Stellar Question	Question favorited by 100 users	3.9k awarded
● Investor	First bounty you offer on another person's question	12.5k awarded
● Nice Question	Question score of 10 or more	338.2k awarded
● Good Question	Question score of 25 or more	100.7k awarded
● Great Question	Question score of 100 or more	15.1k awarded
● Popular Question	Question with 1,000 views	2.6m awarded
● Notable Question	Question with 2,500 views	1.3m awarded
● Famous Question	Question with 10,000 views	325.8k awarded
● Promoter	First bounty you offer on your own question	53.4k awarded
● Scholar	Ask a question and accept an answer	1.2m awarded
● Student	First question with score of 1 or more	1.4m awarded
● Tumbleweed	Asked a question with zero score, no answers, no comments, and low views for a week	698.4k awarded

# What is a question?

- Factoid vs. Broad questions
- In education: Learn something new, or evaluate knowledge
- Could be the goal of your research
- The quality of a question can be equal to the quality of the answer
- Question classification: Categorizing questions
  - WH-words
  - Bag-of-Words and N-grams
  - Word mapping and processing

# Support Vector Machine (SVM)

- Good for regression and classification problems
- Main focus is binary classification
- Often used for text classification
- Separates classes by using a hyperplane
- Four kernels:
  - Linear
  - Radial Basis Function (RBF)
  - Polynomial
  - Sigmoid

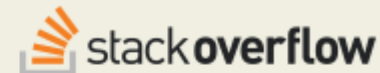


# Methodology

- Data set based on data dump from Stack Exchange Archive
  - Contains XML files based on table content
  - Imported data into MySQL database
  - Imported data from MySQL database into Pandas.DataFrame
- Development: Python 3.5 and Scikit-learn (0.18.dev0)
- Question processing
- Selecting questions and features
- Selecting estimator and parameters for classification

## wordwrap a very long string

Work on work you love. From home.



66

How can you display a long string, website address, word or set of symbols with automatic line breaks to keep a div width? I guess a wordwrap of sorts. Usually adding a space works but is there a CSS solution such as word-wrap?



For example it (very nastily) overlaps divs, forces horizontal scrolling etc.



22

www  
www  
www

What can I add to the above string to fit it neatly within a few lines in a div or within the browser window?

css string long-integer word-wrap

share edit

edited May 13 '09 at 8:57



IAdapter

22.5k 44 131 208

asked May 13 '09 at 6:05



Peter Craig

2,313 14 45 63

5 Even Stack Overflow doesn't word wrap : / – AlbertoPL May 13 '09 at 6:10

14 Well.... 10 months on and it looks like stackoverflow now has a fix for the original messy post - they've added the style word-wrap: break-word; which does the trick in Chrome. – Peter Craig Apr 2 '10 at 9:07

add a comment

# Experiments and Results

- 6 different features
  - Code samples
  - Hexadecimal
  - Homework (synonyms for homework)
  - Links
  - Numerical
  - Tags
  - All features
- 4 different experiments
  - Unprocessed data set vs. all singular feature, and all questions
  - Unprocessed data set vs. all singular features, and question occurrence only
  - Unprocessed data set vs. selected set of features only
  - Stochastic Gradient Descent (SGD) as classifier

# Summary

- Stack overflow as a question quality metric
- Limitations and issues
- Further work
  - Code blocks, Links and Numerical as a feature set
  - Code analysis
  - Sentiment analysis
  - Version numbering

# Demo

# Thanks for listening

# References

- [1] <http://stackoverflow.com/questions/927358/how-do-you-undo-the-last-commit>
- [2] <http://stackoverflow.com/questions/21068763/forcing-a-function-to-return-if-false>
- [3] <http://stackoverflow.com/help/badges>
- [4] <http://stackoverflow.com/questions/856307/wordwrap-a-very-long-string>
- [X]