**Testing core code stability**

**Brief introduction：**

Use fix rate, bug rate and maintain rate to determine the stability of kernel code.

**Variable interpretation:**

Amount: how many records did the committer submit in total

Fix: how many records submitted by the author contain the "fix" field

Bug: how many of all the records submitted by the author contain the "fix" field

Maintain: how many records submitted by the author contain the "fix" field

fix\_ rate：fix\_ rate=fix/amount

bug\_ rate：bug\_ rate=bug/amount

maintain\_ rate：maintain\_ rate=maintain/amount

**Operation outline:**

1. “git log” crawls all the submitted records and writes them to the txt file.

(2) Using Python to sort out the data, which is not UTF-8 encoding data in txt is converted to UTF-8 encoded data.

(3) Calculate four indicators for each contributor: “amount”,” fix”,” bug”and “maintain”.

① Use split to cut out each commit

② Count the commit amount of each author, which is represented by the variable amount.

③ Find out whether there are fix, bug and maintain fields in each commit. If yes, add 1 to the value of each variable, otherwise no operation will be carried out.

(4) Python draws scatter plot, linear regression fitting scatter plot, and calculates the fix rate, bug rate, maintain rate and goodness of fit R ^ 2 of each submitted file.

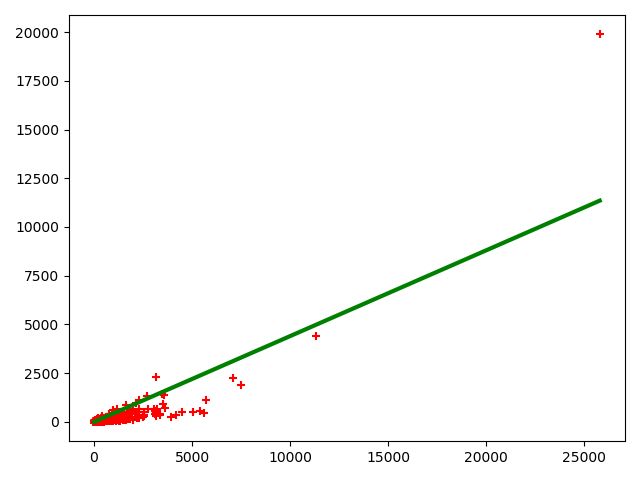
(5) Determine the code stability index, determine the criteria of code stability, that is, fix rate, bug rate, maintain rate.

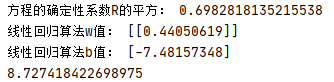
(6) The fix rate, bug rate and maintain rate are compared with the pre-defined indicators. And combined with R ^ 2 to judge whether the code is stable.

**Practical operation:**

1. fix-amount:

After calculation, the fix-amount diagram is as follows:





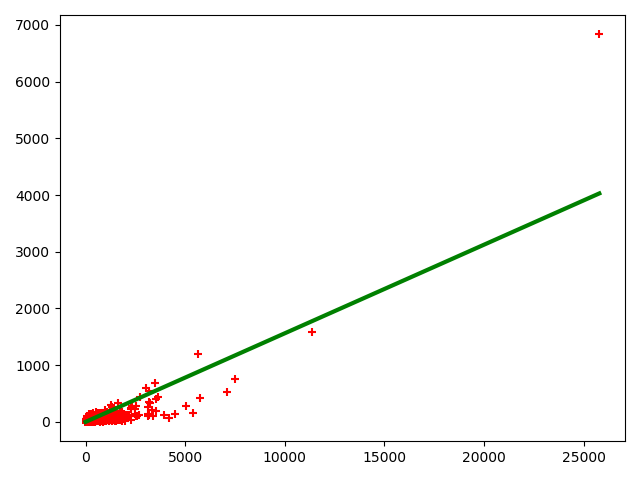
THE PRINT MEANS:

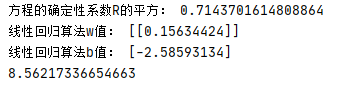
Slope:0.44050619

R^2:0.6982818135215538

1. bug-amount:

After calculation, the bug-amount diagram is as follows:





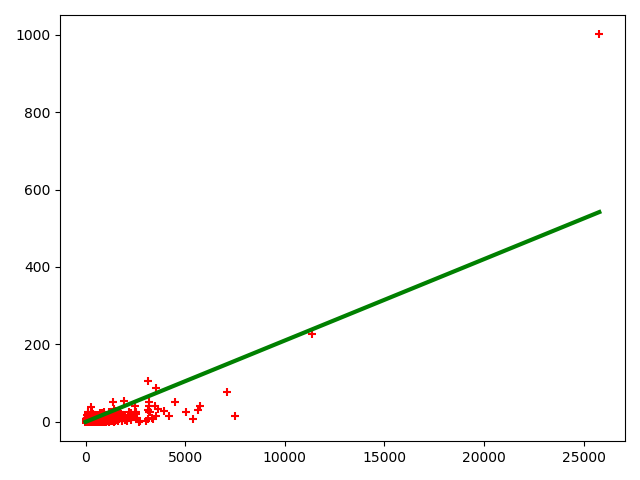
THE PRINT MEANS:

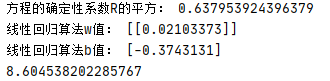
Slope:0.15634424

R^2:0.7143701614808864

1. maintain operation:

After calculation, the maintain-amount diagram is as follows:





THE PRINT MEANS:

Slope:0.02103373

R^2:0.637953924396379

**Evaluation criteria:**

Our criteria are as follows:

* 1. Fix：

Stable：Slope<=0.4

Instable:Slope>0.4

* 1. Bug:

Stable:Slope<=0.2

Instable:Slope>0.2

* 1. Maintain:

Stable:Slope<=0.05

Instable:Slope>0.05

**Comparison and judgment:**

1. From the perspective of R ^ 2, we find that each value of R ^ 2 indicates that the linear regression fits well
2. Comparison and judgmentFix：
   1. Fix:

Test\_Slope=0.44050619>0.4=Standard\_Slope

Conclusion:Instable

* 1. Bug:

Test\_Slope=0.15634424<0.2=Standard\_Slope

Conclusion:Stable

* 1. Maintain:

Test\_Slope=0.02103373<0.05=Standard\_Slope

Conclusion:Stable

**Conclusion:**

Based on the above comparison, we believe that the core code is stable!

**Problems encountered:**

(1) If the file is too large, it is difficult to open, the code will run for a long time and occupy a lot of computer resources.

(2) When Python splits each commit, it can't find the appropriate cutting point. Finally, the author in the second line of each commit is used as the cutting unit.