DStar,MUSE

Org.sfsdaf.sfsdafs.sdfsdf, 139,0.92 T int a;

Org.sfsdaf.sfsdafs.sdfsdf, 139,0.92 F if(a>b)

Org.sfsdaf.sfsdafs.sdfsdf, 139,0.92 T a=10;

Org.sfsdaf.sfsdafs.sdfsdf, 139,0.92 F c=a+b;

Org.sfsdaf.sfsdafs.sdfsdf, 139,0.91

Org.sfsdaf.sfsdafs.sdfsdf, 139,0.90

Org.sfsdaf.sfsdafs.sdfsdf, 139,0.89

Org.sfsdaf.sfsdafs.sdfsdf, 139,0.88

Org.sfsdaf.sfsdafs.sdfsdf, 139,0.80

定特征，有哪些。

1、a，AST树分析，整个程序的，遍历到语句，提取其中的特征

1、b，当做文本，提取特征

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 变量数 | 符号数 | 。。。 |  | 标签 | 预测标签 |
|  |  |  |  | T | T（TP） |
|  |  |  |  | F | F（TN） |
|  |  |  |  | T | F（FN） |
|  |  |  |  | F | T（FP） |

395个版本，并列第一的语句，5000条记录。

分类器90%训练，10预测，

评价1，F1、precision评价

评价2，exam，inspect

版本A，10条并列第一，8条是真有问题。Inspect到第8条才能发现缺陷

10条，分类，有问题3条，没问题7条，inspect第二条就发现，

Einspect@n（1,2,3,5，10）

EXAM

Einspect@1，395个版本，检查第一条，60个版本发现

Einspect@3，395个版本，检查第一条，90个版本发现

EXAM，1个版本，9000行，第八行发现，8/9000