Advance Software Engineering (SE-406)

LAB A1-G3

Laboratory Manual



Department of Software Engineering

DELHI TECHNOLOGICAL UNIVERSITY(DTU)

Shahbad Daulatpur, Bawana Road, Delhi-110042

Submitted to: -

Submitted by:-

Ms. Parul Sharma

Name: ASHISH KUMAR

Roll number: 2K18/SE/041

EXPERIMENT 5

- ASHISH KUMAR
- 2K18/SE/041

Aim: Data Acquisition for software defect prediction.

About dataset:-

Link to dataset: http://promise.site.uottawa.ca/SERepository/datasets/cm1.arff

This is a PROMISE data set made publicly available in order to encourage repeatable, verifiable, refutable, and/or improvable predictive models of software engineering.

- 1. Sources:
 - -- Creators: NASA, then the NASA Metrics Data Program,
 - -- http://mdp.ivv.nasa.gov. Contacts: Mike Chapman, Galaxy Global Corporation
 - -- Donor: Tim Menzies (tim@barmag.net)
 - -- Date: December 2 2004
- 2. Number of instances: 10885
- 3. Number of attributes: 22 (5 different lines of code measure, 3 McCabe metrics, 4 base Halstead measures, 8 derived Halstead measures, a branch-count, and 1 goal field)
- 4. Attribute Information:

```
1. loc
            : numeric % McCabe's line count of code
2. v(g)
            : numeric % McCabe "cyclomatic complexity"
            : numeric % McCabe "essential complexity"
3. \operatorname{ev}(g)
4. iv(g)
            : numeric % McCabe "design complexity"
            : numeric % Halstead total operators + operands
5. n
            : numeric % Halstead "volume"
6. v
7.1
            : numeric % Halstead "program length"
            : numeric % Halstead "difficulty"
8. d
            : numeric % Halstead "intelligence"
9. i
10. e
            : numeric % Halstead "effort"
11. b
            : numeric % Halstead
            : numeric % Halstead's time estimator
12. t
13. lOCode: numeric % Halstead's line count
                   : numeric % Halstead's count of lines of comments
14. lOComment
```

- 15. IOBlank : numeric % Halstead's count of fines of comment
- 16. lOCodeAndComment: numeric

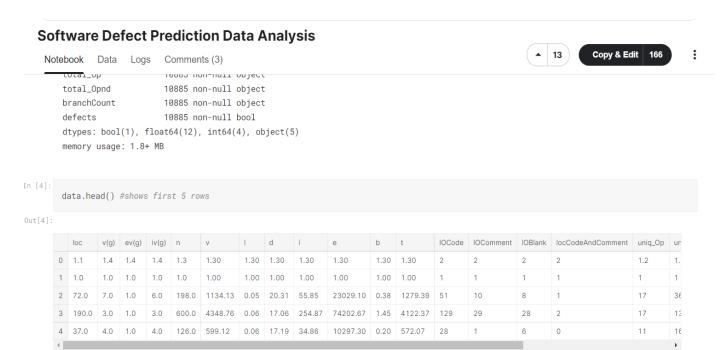
17. uniq_Op : numeric % unique operators
18. uniq_Opnd : numeric % unique operands
19. total_Op : numeric % total operators
20. total_Opnd : numeric % total operands
21: branchCount : numeric % of the flow graph

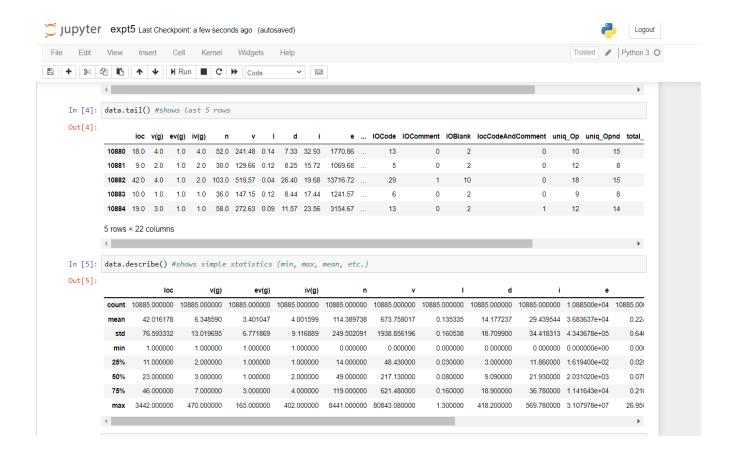
22. defects : {false,true} % module has/has not one or more reported defects

5. Missing attributes: none

6. Class Distribution: the class value (defects) is discrete

false: 2106 = 19.35% true: 8779 = 80.65%





Learning from experiment: We are successful in finding a defect dataset and we analyzed it properly.