# Report

## Practical assignment

## Exploratory analysis of the data

## The single tree

PreRel defects <= 4

PreRel defects > 4

NOT\_avg <= 0.84167

NOT\_avg > 0.84167

VG\_max > 26

VG\_max <= 26

## Values

### Analysis 1 (single tree)

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | **Defects are observed** | |
|  |  | True | False |
| **Model predicts** | Positive | 229 | 93 |
| **defect.** | Negative | 84 | 255 |

|  |  |  |
| --- | --- | --- |
| Accuracy | Precision | Recall |
| 0.7322 | 0.7111 | 0,7316 |

### Analysis 2 (bagging)

### Analysis 3 (random forests)

## Discussion of the differences

### Single Tree vs Bagging

### Bagging vs Random Forests

McNemar's chi-squared = 2.1607, df = 1, p-value = 0.1416

## Note

## Appendix 1: Metrics in the Eclipse Data Set