Assignment 1 Team 3

# Most frequent terms in the corpus with tf

### FreeMind

1. ('node', 9463)
2. ('map', 6835)
3. ('string', 5980)
4. ('action', 5600)
5. ('set', 4839)
6. ('control', 4399)
7. ('mind', 3717)
8. ('view', 2890)
9. ('select', 2745)
10. ('name', 2710)

### Weka

1. ('set', 28090)
2. ('string', 27302)
3. ('instanc', 26212)
4. ('valu', 24211)
5. ('option', 18234)
6. ('attribut', 17423)
7. ('return', 16781)
8. ('class', 16447)
9. ('result', 15788)
10. ('name', 15548)

# Highest document frequency

### FreeMind

1. ('free', 489)
2. ('modifi', 487)
3. ('thi', 487)
4. ('version', 487)
5. ('write', 487)
6. ('copyright', 487)
7. ('public', 486)
8. ('softwar', 486)
9. ('purpos', 485)
10. ('later', 485)

### Weka

1. ('thi', 1640)
2. ('ani', 1639)
3. ('http', 1638)
4. ('gener', 1637)
5. ('www', 1630)
6. ('org', 1626)
7. ('java', 1620)
8. ('new', 1467)
9. ('distribut', 1390)
10. ('version', 1385)

# Documents with highest number of unique terms

### FreeMind

1. ('freemind/freemind/modes/mindmapmode/MindMapController.java', 9229)
2. ('freemind/plugins/map/FreeMindMapController.java', 9082)
3. ('freemind/freemind/controller/Controller.java', 7431)

### Weka

1. ('weka-3-7-13/src/main/java/weka/gui/beans/KnowledgeFlowApp.java', 25151)
2. ('weka-3-7-13/src/main/java/weka/classifiers/bayes/net/GUI.java', 13246)
3. ('weka-3-7-13/src/main/java/weka/classifiers/evaluation/Evaluation.java', 12987)

# Documents with the least amount of unique terms

### FreeMind

1. ('freemind/freemind/modes/IconInformation.java', 22)
2. ('freemind/freemind/modes/common/plugins/NodeNoteBase.java', 77)
3. ('freemind/freemind/modes/mindmapmode/actions/HookAction.java', 79)

### Weka

1. ('weka-3-7-13/src/main/java/weka/core/json/sym.java', 33)
2. ('weka-3-7-13/src/main/java/weka/core/expressionlanguage/parser/sym.java', 38)
3. ('weka-3-7-13/src/main/java/weka/core/pmml/jaxbbindings/package-info.java', 42)