**INSTITUTO FEDERAL DO MARANHÃO-IFMA**

**CAMPUS SANTA INÊS**

**TURMA: ENG. DA COMPUTAÇÃO - 2022**

**DISCIPLINA: INTRO. À ENGENHARIA DA COMPUTAÇÃO**

**PROF. EMANUEL CLEYTON MACEDO LEMOS**

**ALUNO(A): FÁBIO DA SILVA ELOI JÚNIOR**

**RELATÓRIO DE DADOS – 2**

**WEKA: trees.J48 e rules.DecisionTable.**

**CONCLUSÕES**

Na base de dados “car”, utilizei as técnicas de mineração: trees.J48 e rules.DecisionTable. Ambas as técnicas de mineração de dados trabalharam com 1728 instâncias e os 7 atributos. Algumas diferenças que podem ser citadas entre elas é que uma utiliza um número de regras para a análise e a outra trabalha utilizando número de folhas da árvore e o tamanho da árvore de decisão. O tempo necessário para agrupar e concluir o modelo é diferente em ambas as classificações, onde a trees.J48 consegue fazer com o tempo de 0.04 segundos, já a rules.DecisionTable apresenta seu modelo em 0.12 segundos. O tempo necessário para testar o modelo em dados de treinamento também é diferente, porém, dessa vez classificação DecisionTable gastou 0.03 segundos para testar, e a J48 precisou de 0.08 segundos.

Houve também uma divergência na quantidade de instancias classificadas de forma correta, onde essa diferença trouxe diversas mudanças entre elas. O erro absoluto relativo e o erro quadrático relativo de raiz na rules.DecisionTable obteve uma porcentagem muito maior com relação a outra classificação analisada. Outra questão que diverge entre elas é a precisão detalhada por classe, onde cada uma apresenta diferentes valores para a classes da análise dos dados feita, e também uma diferença na matriz de confusão de ambas, onde a J48 consegue apresentar com mais detalhes os valores.

**Resultado da Classificação com o J48:**

**=== Run information ===**

**Scheme: weka.classifiers.trees.J48 -C 0.25 -M 2**

**Relation: car**

**Instances: 1728**

**Attributes: 7**

**buying**

**maint**

**doors**

**persons**

**lug\_boot**

**safety**

**class**

**Test mode: evaluate on training data**

**=== Classifier model (full training set) ===**

**J48 pruned tree**

**------------------**

**safety = low: unacc (576.0)**

**safety = med**

**| persons = 2: unacc (192.0)**

**| persons = 4**

**| | buying = vhigh**

**| | | maint = vhigh: unacc (12.0)**

**| | | maint = high: unacc (12.0)**

**| | | maint = med**

**| | | | lug\_boot = small: unacc (4.0)**

**| | | | lug\_boot = med: unacc (4.0/2.0)**

**| | | | lug\_boot = big: acc (4.0)**

**| | | maint = low**

**| | | | lug\_boot = small: unacc (4.0)**

**| | | | lug\_boot = med: unacc (4.0/2.0)**

**| | | | lug\_boot = big: acc (4.0)**

**| | buying = high**

**| | | lug\_boot = small: unacc (16.0)**

**| | | lug\_boot = med**

**| | | | doors = 2: unacc (4.0)**

**| | | | doors = 3: unacc (4.0)**

**| | | | doors = 4: acc (4.0/1.0)**

**| | | | doors = 5more: acc (4.0/1.0)**

**| | | lug\_boot = big**

**| | | | maint = vhigh: unacc (4.0)**

**| | | | maint = high: acc (4.0)**

**| | | | maint = med: acc (4.0)**

**| | | | maint = low: acc (4.0)**

**| | buying = med**

**| | | maint = vhigh**

**| | | | lug\_boot = small: unacc (4.0)**

**| | | | lug\_boot = med: unacc (4.0/2.0)**

**| | | | lug\_boot = big: acc (4.0)**

**| | | maint = high**

**| | | | lug\_boot = small: unacc (4.0)**

**| | | | lug\_boot = med: unacc (4.0/2.0)**

**| | | | lug\_boot = big: acc (4.0)**

**| | | maint = med: acc (12.0)**

**| | | maint = low**

**| | | | lug\_boot = small: acc (4.0)**

**| | | | lug\_boot = med: acc (4.0/2.0)**

**| | | | lug\_boot = big: good (4.0)**

**| | buying = low**

**| | | maint = vhigh**

**| | | | lug\_boot = small: unacc (4.0)**

**| | | | lug\_boot = med: unacc (4.0/2.0)**

**| | | | lug\_boot = big: acc (4.0)**

**| | | maint = high: acc (12.0)**

**| | | maint = med**

**| | | | lug\_boot = small: acc (4.0)**

**| | | | lug\_boot = med: acc (4.0/2.0)**

**| | | | lug\_boot = big: good (4.0)**

**| | | maint = low**

**| | | | lug\_boot = small: acc (4.0)**

**| | | | lug\_boot = med: acc (4.0/2.0)**

**| | | | lug\_boot = big: good (4.0)**

**| persons = more**

**| | lug\_boot = small**

**| | | buying = vhigh: unacc (16.0)**

**| | | buying = high: unacc (16.0)**

**| | | buying = med**

**| | | | maint = vhigh: unacc (4.0)**

**| | | | maint = high: unacc (4.0)**

**| | | | maint = med: acc (4.0/1.0)**

**| | | | maint = low: acc (4.0/1.0)**

**| | | buying = low**

**| | | | maint = vhigh: unacc (4.0)**

**| | | | maint = high: acc (4.0/1.0)**

**| | | | maint = med: acc (4.0/1.0)**

**| | | | maint = low: acc (4.0/1.0)**

**| | lug\_boot = med**

**| | | buying = vhigh**

**| | | | maint = vhigh: unacc (4.0)**

**| | | | maint = high: unacc (4.0)**

**| | | | maint = med: acc (4.0/1.0)**

**| | | | maint = low: acc (4.0/1.0)**

**| | | buying = high**

**| | | | maint = vhigh: unacc (4.0)**

**| | | | maint = high: acc (4.0/1.0)**

**| | | | maint = med: acc (4.0/1.0)**

**| | | | maint = low: acc (4.0/1.0)**

**| | | buying = med: acc (16.0/5.0)**

**| | | buying = low**

**| | | | maint = vhigh: acc (4.0/1.0)**

**| | | | maint = high: acc (4.0)**

**| | | | maint = med: good (4.0/1.0)**

**| | | | maint = low: good (4.0/1.0)**

**| | lug\_boot = big**

**| | | buying = vhigh**

**| | | | maint = vhigh: unacc (4.0)**

**| | | | maint = high: unacc (4.0)**

**| | | | maint = med: acc (4.0)**

**| | | | maint = low: acc (4.0)**

**| | | buying = high**

**| | | | maint = vhigh: unacc (4.0)**

**| | | | maint = high: acc (4.0)**

**| | | | maint = med: acc (4.0)**

**| | | | maint = low: acc (4.0)**

**| | | buying = med**

**| | | | maint = vhigh: acc (4.0)**

**| | | | maint = high: acc (4.0)**

**| | | | maint = med: acc (4.0)**

**| | | | maint = low: good (4.0)**

**| | | buying = low**

**| | | | maint = vhigh: acc (4.0)**

**| | | | maint = high: acc (4.0)**

**| | | | maint = med: good (4.0)**

**| | | | maint = low: good (4.0)**

**safety = high**

**| persons = 2: unacc (192.0)**

**| persons = 4**

**| | buying = vhigh**

**| | | maint = vhigh: unacc (12.0)**

**| | | maint = high: unacc (12.0)**

**| | | maint = med: acc (12.0)**

**| | | maint = low: acc (12.0)**

**| | buying = high**

**| | | maint = vhigh: unacc (12.0)**

**| | | maint = high: acc (12.0)**

**| | | maint = med: acc (12.0)**

**| | | maint = low: acc (12.0)**

**| | buying = med**

**| | | maint = vhigh: acc (12.0)**

**| | | maint = high: acc (12.0)**

**| | | maint = med**

**| | | | lug\_boot = small: acc (4.0)**

**| | | | lug\_boot = med: acc (4.0/2.0)**

**| | | | lug\_boot = big: vgood (4.0)**

**| | | maint = low**

**| | | | lug\_boot = small: good (4.0)**

**| | | | lug\_boot = med: good (4.0/2.0)**

**| | | | lug\_boot = big: vgood (4.0)**

**| | buying = low**

**| | | maint = vhigh: acc (12.0)**

**| | | maint = high**

**| | | | lug\_boot = small: acc (4.0)**

**| | | | lug\_boot = med: acc (4.0/2.0)**

**| | | | lug\_boot = big: vgood (4.0)**

**| | | maint = med**

**| | | | lug\_boot = small: good (4.0)**

**| | | | lug\_boot = med: good (4.0/2.0)**

**| | | | lug\_boot = big: vgood (4.0)**

**| | | maint = low**

**| | | | lug\_boot = small: good (4.0)**

**| | | | lug\_boot = med: good (4.0/2.0)**

**| | | | lug\_boot = big: vgood (4.0)**

**| persons = more**

**| | buying = vhigh**

**| | | maint = vhigh: unacc (12.0)**

**| | | maint = high: unacc (12.0)**

**| | | maint = med: acc (12.0/1.0)**

**| | | maint = low: acc (12.0/1.0)**

**| | buying = high**

**| | | maint = vhigh: unacc (12.0)**

**| | | maint = high: acc (12.0/1.0)**

**| | | maint = med: acc (12.0/1.0)**

**| | | maint = low: acc (12.0/1.0)**

**| | buying = med**

**| | | maint = vhigh: acc (12.0/1.0)**

**| | | maint = high: acc (12.0/1.0)**

**| | | maint = med**

**| | | | lug\_boot = small: acc (4.0/1.0)**

**| | | | lug\_boot = med: vgood (4.0/1.0)**

**| | | | lug\_boot = big: vgood (4.0)**

**| | | maint = low**

**| | | | lug\_boot = small: good (4.0/1.0)**

**| | | | lug\_boot = med: vgood (4.0/1.0)**

**| | | | lug\_boot = big: vgood (4.0)**

**| | buying = low**

**| | | maint = vhigh: acc (12.0/1.0)**

**| | | maint = high**

**| | | | lug\_boot = small: acc (4.0/1.0)**

**| | | | lug\_boot = med: vgood (4.0/1.0)**

**| | | | lug\_boot = big: vgood (4.0)**

**| | | maint = med**

**| | | | lug\_boot = small: good (4.0/1.0)**

**| | | | lug\_boot = med: vgood (4.0/1.0)**

**| | | | lug\_boot = big: vgood (4.0)**

**| | | maint = low**

**| | | | lug\_boot = small: good (4.0/1.0)**

**| | | | lug\_boot = med: vgood (4.0/1.0)**

**| | | | lug\_boot = big: vgood (4.0)**

**Number of Leaves : 131**

**Size of the tree : 182**

**Time taken to build model: 0.04 seconds**

**=== Evaluation on training set ===**

**Time taken to test model on training data: 0.08 seconds**

**=== Summary ===**

**Correctly Classified Instances 1664 96.2963 %**

**Incorrectly Classified Instances 64 3.7037 %**

**Kappa statistic 0.9198**

**Mean absolute error 0.0248**

**Root mean squared error 0.1114**

**Relative absolute error 10.8411 %**

**Root relative squared error 32.9501 %**

**Total Number of Instances 1728**

**=== Detailed Accuracy By Class ===**

**TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class**

**0,977 0,019 0,992 0,977 0,984 0,948 0,997 0,998 unacc**

**0,964 0,028 0,907 0,964 0,934 0,916 0,996 0,978 acc**

**0,826 0,007 0,838 0,826 0,832 0,825 0,997 0,916 good**

**0,846 0,003 0,917 0,846 0,880 0,876 0,999 0,952 vgood**

**Weighted Avg. 0,963 0,020 0,964 0,963 0,963 0,933 0,997 0,989**

**=== Confusion Matrix ===**

**a b c d <-- classified as**

**1182 25 3 0 | a = unacc**

**10 370 2 2 | b = acc**

**0 9 57 3 | c = good**

**0 4 6 55 | d = vgood**

**Resultado da Classificação com DecisionTable**

**=== Run information ===**

**Scheme: weka.classifiers.rules.DecisionTable -X 1 -S "weka.attributeSelection.BestFirst -D 1 -N 5"**

**Relation: car**

**Instances: 1728**

**Attributes: 7**

**buying**

**maint**

**doors**

**persons**

**lug\_boot**

**safety**

**class**

**Test mode: evaluate on training data**

**=== Classifier model (full training set) ===**

**Decision Table:**

**Number of training instances: 1728**

**Number of Rules : 432**

**Non matches covered by Majority class.**

**Best first.**

**Start set: no attributes**

**Search direction: forward**

**Stale search after 5 node expansions**

**Total number of subsets evaluated: 22**

**Merit of best subset found: 94.329**

**Evaluation (for feature selection): CV (leave one out)**

**Feature set: 1,2,4,5,6,7**

**Time taken to build model: 0.12 seconds**

**=== Evaluation on training set ===**

**Time taken to test model on training data: 0.03 seconds**

**=== Summary ===**

**Correctly Classified Instances 1662 96.1806 %**

**Incorrectly Classified Instances 66 3.8194 %**

**Kappa statistic 0.9189**

**Mean absolute error 0.2225**

**Root mean squared error 0.262**

**Relative absolute error 97.1571 %**

**Root relative squared error 77.4937 %**

**Total Number of Instances 1728**

**=== Detailed Accuracy By Class ===**

**TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class**

**0,965 0,000 1,000 0,965 0,982 0,945 0,989 0,993 unacc**

**0,961 0,029 0,904 0,961 0,932 0,912 0,984 0,927 acc**

**0,870 0,007 0,833 0,870 0,851 0,845 0,975 0,847 good**

**1,000 0,009 0,813 1,000 0,897 0,897 0,998 0,901 vgood**

**Weighted Avg. 0,962 0,007 0,965 0,962 0,963 0,932 0,987 0,969**

**=== Confusion Matrix ===**

**a b c d <-- classified as**

**1168 39 3 0 | a = unacc**

**0 369 9 6 | b = acc**

**0 0 60 9 | c = good**

**0 0 0 65 | d = vgood**