A1

a)

Anzahl der Regeln

|  |  |  |  |
| --- | --- | --- | --- |
|  | Contact Lenses | Cars | Kr vs Kp |
| Jrip | 2 Regeln | Einige Regeln mit guter generalisierung | 15 Regeln |
| Conjunctive | Nur default Rule | Nur default Rule | Eine Regel |
| Prism | Sehr viele Regeln (overfitted) | Sehr viele Regeln (of) | Sehr viele Regeln (of) |

Bedingungen pro Regel

|  |  |  |  |
| --- | --- | --- | --- |
|  | Contact Lenses | Cars | Kr vs Kp |
| Jrip | 1-2 | 3-4 | 2-3 |
| Conjunctive | 0 | 0 | 1 |
| Prism | Ca 3 | Ca 5 | Ca 4 |

Vorhergesagte Klassen

|  |  |  |  |
| --- | --- | --- | --- |
|  | Contact Lenses | Cars | Kr vs Kp |
| Jrip | 3 Klassen 88% richtig | 4 Klassen 94 % richtig | 2 Klassen 99,5% richtig |
| Conjunctive | 3 Klassen 63% richtig | 4 Klassen 70% richtig | 2 Klassen 66% richtig |
| Prism | 3 Klassen 100% richtig | 4 Klassen 100% richtig | 2 Klassen 100% richtig |

Conjunctive Rule: Versucht alles mit einer Regel abzudecken (allgemeinste lösung)

Prism: Versucht für jeden Datenpunkt eine Regel zu erstellen (overfitting)

Jrip: Mittelweg der eine generalisierte Lösung findet ohne zu viele Datenpunkte falsch zu klassifizieren und ohne zu viel overfitting zu erzeugen

b)

Default Rule:

|  |  |  |  |
| --- | --- | --- | --- |
|  | Contact Lenses | Cars | Kr vs Kp |
| Jrip | contact-lenses=none (12.0/0.0) | class=unacc (1164.0/7.0) | class=won (1667.0/6.0) |
| Conjunctive | contact-lenses = none | class = unacc | class = nowin |
| Prism | - | - | - |

Jrip:

Conjunctive: meistens wird die häufigste Klasse als default Rule ausgewählt ? / besser als nur eine Klassifikation

Prism: keine default Rule / keine Güte, da nicht abgedeckte Daten als unklassifiziert eingestuft werden

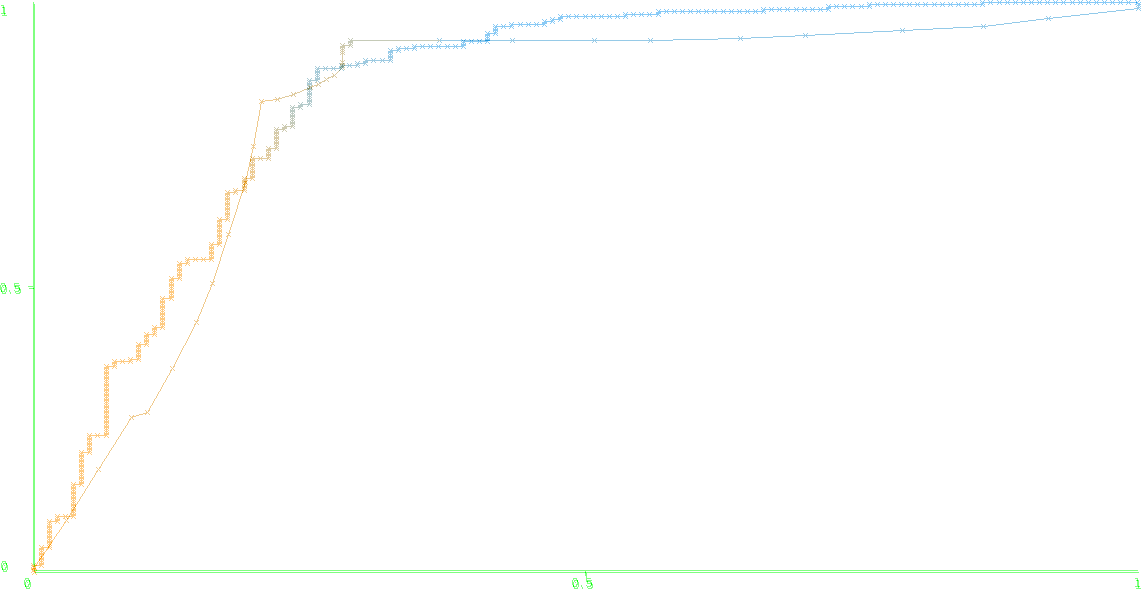
Jrip: meiste Attribute? Beste güte, da die Default Rule so gewählt wird, dass der fehler möglichst gering wird

c) Die Datenmengen mit den wenigsten Daten sind am einfachsten zu lernen (Contact Lenses) allerdings ergeben die Datensätze mit mehr Daten (kr vs kp) bessere Regeln.

d) Jrip ist allgemeiner, da hier nicht versucht wird Daten auswendig zu lernen (weniger overfitting), sondern durch eine generelle Regel auch mit nicht im Datensatz vorhandenen Daten umzugehen. Prism deckt alle Daten aus der Trainingsmenge mit Regeln ab und kann deshalb nicht mit unbekannten Daten umgehen, also nicht so gut dazulernen.

A3

ROC Kurven für colic.arff für NaiveBayes und J48



1. - Die Fläche unter der NaiveBayes Kurve ist etwas größer als die unter der J48.

- NaiveBayes hat mehr Datenpunkte als J48.

- Die ROC Kurve von J48 ist

b) –Der NaiveBayes Algorithmus ist besser auf dem colic Datensatz, da die Fläche under der Kurve (AUC) größer ist und so der Anteil der True Positives gegenüber den False Positives höher ist, was bedeutet, dass die Daten besser Klassifiziert wurden.