

Aluno: Railson Martins da Mata

1)

- Robustos para pontos de ruído isolados (apresentam pouco impacto na probabilidade)
- Podem tratar valores faltantes (ignorando-os)
- Robustos para lidar com atributos irrelevantes (a probabilidade se torna quase uniformemente distribuída)
- Atributos correlacionados podem degradar o desempenho porque a suposição de
- independência condicional não é mais válida para tais atributos

3) Eu não consegui realizar a implementação do algoritmo Bayes Simples de forma genérica, onde a implementação que eu fiz, foi para o caso específico da tabela que define quem é ou não inadimplente, tal algoritmo implementado se mostrou extremamente eficiente, tendo acertado todas as classificações. Eu realizei alguns testes, usando algoritmos prontos da ferramenta Weka, tanto para a base de dados iris, quando para a base de dados de inadimplência. Os resultados podem ser vistos a seguir:

Para a Base de dados Inadimplência:

- Na ferramenta Weka o algoritmo *Naive Bayes Simple* para a base de dados inadimplência, acertou todas as classificações usando a opção de conjunto de treinamento. O mesmo algoritmo com o Cross-validation com 10 folds, acertou apenas 90% das classificações.
- O algoritmo *Naive bayes* usando conjunto de treinamento, acertou também todas as classificações, para a base de dados inadimplência, entretanto, o mesmo algoritmo usando a opção de Cross-validation acertou apenas 70% das classificações.
- O algoritmo *Naive Bayes Updatable* teve o mesmo desempenho do item anterior.
- Tanto os algoritmos do Naive Bayes: *Multinomial e Multinomial Updatable*, não funcionaram para essa base de dados.

Para a Base de dados Iris:

- O algoritmo *Bayes Net* acertou 94.66% das classificações usando a opção conjunto de treinamento, já com a opção Cross-Validation 10 folds, ele acertou 92.66% das classificações.
- O algoritmo *Naive Bayes* acertou 96% das classificações usando a opção conjunto de treinamento, já com a opção Cross-Validation 10 folds, ele acertou também 96% das classificações.

- O algoritmo *Naive Bayes Multinomial* acertou 95.33% das classificações usando a opção conjunto de treinamento, já com a opção Cross-Validation 10 folds, ele acertou também 95.33% das classificações.
- O algoritmo *Naive Bayes Simple* acertou 96% das classificações usando a opção conjunto de treinamento, já com a opção Cross-Validation 10 folds, ele acertou apenas 95.33% das classificações.
- O algoritmo *Naive Bayes Updatable* acertou 96% das classificações usando a opção conjunto de treinamento, já com a opção Cross-Validation 10 folds, ele acertou também 96% das classificações.

Como visto para a base de dados inadimplência o Naive Bayes Simple foi o que se mostrou mais eficiente na realização das classificações. Já para a base de dados Iris o que se mostrou melhor foi o Naive Bayes, com acertos de 96% nas duas opções selecionadas.