



Instituto Federal de Educação, Ciência e Tecnologia do Ceará
Programa de Pós-Graduação em Ciência da Computação
Departamento de Telemática

Relatório de experimento

Trabalho 4
Classificador Bayesiano com opção de rejeição

Alan Rabelo Martins

2019

1 Introdução

Neste trabalho apresento os resultados do classificador Bayesiano usando a função de probabilidade Gaussiana pura com a opção de Rejeição. Na seção 2 apresento a metodologia e na 3 apresento os resultados alcançados.

2 Metodologia

Os algoritmos foram desenvolvidos em Python 3.6 utilizando a IDE Pycharm e apenas a biblioteca numpy e a matplotlib. Para avaliação do classificador Bayesiano foram utilizados 5 folds com 30 rodadas de teste e computadas a acurácia e o desvio padrão. O parâmetro W_r foi variado de 0.04 até 0.48

3 Resultados do Classificador

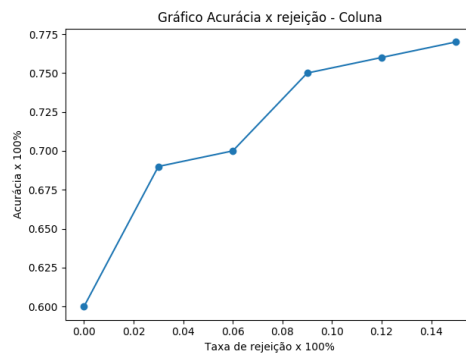
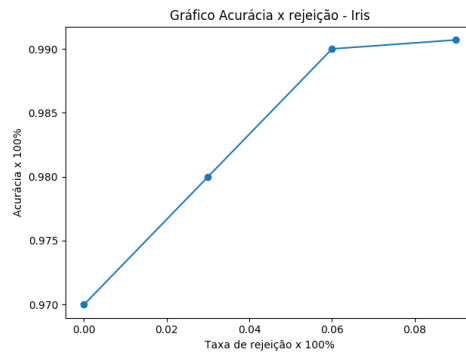
Após aplicar a metodologia nos datasets podemos verificar o resultado do experimento em termos de acurácia e desvio padrão exibidos nas seções abaixo. A avaliação se dá pelo valor de acurácia dos modelos e seu respectivo desvio padrão além da taxa de rejeição e seu desvio padrão através das iterações.

Para efeito de comparação no dataset da Íris os valores de acurácia ficaram em 99.07% após a rejeição de 0.09% dos dados de teste. Este resultado foi encontrado com o $W_r = 0.12$. O desvio padrão das taxas de acerto foi 0.02 enquanto que da taxa de rejeição foi de 0.04.

No dataset da Coluna vertebral, os valores de acurácia ficaram em 77.6% após a rejeição de 5% dos dados de teste. Este resultado foi encontrado com o $W_r = 0.24$. O desvio padrão das taxas de acerto foi 0.08 enquanto que da taxa de rejeição foi de 0.01.

No dataset artificial, os valores de acurácia ficaram em 100% após a rejeição de 0.03% dos dados de teste. Este resultado foi encontrado com o $W_r = 0.04$. O desvio padrão das taxas de acerto foi 0.00 enquanto que da taxa de rejeição foi de 0.04.

Abaixo podemos verificar os gráficos que relacionam a taxa de acurácia com a rejeição em cada dataset.



4 Conclusão

A partir dos resultados podemos constatar que o algoritmo gera bons resultados apesar da acurácia mais baixa encontrada nos dados da coluna vertebral. No entanto em certos casos, especialmente no diagnóstico de doenças pode ser imprescindível dar um resultado com uma certeza maior, descartando qualquer possibilidade de erro.